

Université de Sherbrooke

**Investigation des mécanismes de régulations post-traductionnelles  
médiées par l'ubiquitine et par un nouveau variant : Ub<sup>KEKS</sup>**

Par  
Marie-Line Dubois  
Programme de biologie cellulaire

Thèse présentée à la Faculté de médecine et des sciences de la santé  
en vue de l'obtention du grade de philosophiae doctor (Ph. D.)  
en biologie cellulaire

Sherbrooke, Québec, Canada  
Septembre 2019

Membres du jury d'évaluation  
Pr François-Michel Boisvert Ph.D, Département d'immunologie et biologie cellulaire  
Pr Steve Jean Ph.D, Département d'immunologie et biologie cellulaire  
Pr Jean-Luc Parent Ph.D, Département de médecine  
Pre Amélie Fradet-Turcotte Ph.D, Département de Biologie Moléculaire, de  
Biochimie Médicale et de Pathologie, Université Laval

© [Marie-Line Dubois, 2019]

*Parce qu'il est un modèle de persévérance et de courage et  
parce qu'il fait preuve d'une force de caractère étonnante  
Je dédie ma thèse de doctorat à mon conjoint Sofiane*

*Ce n'est pas le vent qui décide de la destination,  
c'est l'orientation que tu donnes à ta voile.  
Le vent est le même pour tous.*

*-Jim Rohn-*

## Remerciements

*J'aimerais tout d'abord remercier les membres de mon jury d'avoir pris de leur temps pour évaluer ma thèse. Merci à mon comité d'encadrement, Pr Jean-Luc Parent et Pre Caroline Saucier d'avoir été présent tout au long de mon parcours.*

*J'aimerais particulièrement remercier mon directeur de thèse François-Michel Boisvert, pour tout ton temps, toute la patience dont tu as fait preuve envers moi et mes lubies scientifiques. Pour les opportunités que tu m'as offertes et pour toute la confiance que tu as placée en moi.*

*Un merci particulier au Pr.Xavier Roucou et aux membres de son équipe qui ont été les initiateurs du projet UBBP4 mais qui m'ont fait confiance pour le mener à bien. Je suis très reconnaissante de toute l'aide qu'ils m'ont apporté dans la réalisation de ce projet.*

*Je tiens à dire un énorme merci à Dominique Lévesque pour son aide dans le laboratoire autant au niveau scientifique que personnel. Tu as été un collègue et un ami précieux durant mes années au laboratoire.*

*À Anaïs, Alyson, Élie, Jennif, Chopchop et Pat, un merci sincère d'avoir mis de la joie dans mes journées lors de ma maîtrise et de mon doctorat. À toutes les différentes générations d'étudiants que j'ai vu défiler dans le laboratoire, un gros merci pour avoir fait de mon temps dans le lab un moment un peu moins pénible.*

*À mon trio de femmes FORTES, INDÉPENDANTES et qui RÉUSSISENT ce qu'elles entreprennent (Jenny, Marie et Céline). Merci d'avoir été des amies si formidables qui m'ont accompagné ces dernières années. D'avoir été là dans des moments difficiles et d'avoir créé des moments si magiques par votre présence dans ma vie.*

*À mon conjoint Sofiane, ta patience, ton amour et ta confiance inébranlable en mes capacités me redonnent confiance quand je suis certaine que je ne pourrais pas me rendre jusqu'au bout. Je me considère choyé de pouvoir t'avoir avec moi à mes côtés.*

*À ma famille, Guylaine, Jacques, Jérôme et Alex, je vous remercie du fond du cœur pour votre soutien, votre amour inconditionnel et d'avoir toujours été présents. Je vous aime de tout mon cœur.*

*Finalement, je remercie tous ceux qui ont contribué de près ou de loin à la réussite de ce travail.*

## Résumé

### Investigation des mécanismes de régulations post-traductionnelles médiées par l'ubiquitine et par un nouveau variant : Ub<sup>KEKS</sup>

Par

Marie-Line Dubois

Programme de biologie cellulaire

Thèse présentée à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de philosophiae doctor (Ph. D.) en biologie cellulaire, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

L'ubiquitinylation est une modification post-traductionnelle qui peut moduler la localisation, les interactions ou encore la stabilité d'une protéine. Cette dernière est impliquée dans presque tous les processus cellulaires tels que la réplication et la réparation de l'ADN, dans le contrôle du cycle cellulaire, dans la réponse du système immunitaire. Pour s'attacher de manière covalente à son substrat, cette dernière a besoin d'une machinerie enzymatique qui implique les protéines E1, E2 et E3. Ce sont les protéines de la classe des E3 qui sont responsables de la reconnaissance des substrats. Les protéines DDB1-Cul4 forment l'un des complexes de cette famille responsable de l'ubiquitinylation de nombreuses protéines impliquées dans divers mécanismes cellulaires. Pour cibler ses différents substrats, il utilise des protéines de reconnaissance nommées DCAF. Cette famille de protéines se compose de près de 60 membres qui ont été caractérisés pour la plupart par des études à large échelle de protéomique et de bio-informatique. Dans la première partie de cette thèse, nous avons utilisé différentes approches de protéomique afin de valider l'appartenance de ces protéines à cette famille et aussi d'identifier par une seconde approche les cibles de chacune de celles-ci.

La seconde partie de ce travail porte sur la caractérisation du gène UBBP4 qui a initialement été annoté comme pseudogène de l'ubiquitine. Cependant, grâce à la découverte de petits cadres de lecture alternatif (altORF) et à la réanalyse de données de spectrométrie de masse, nous avons identifié 3 variants de l'ubiquitine provenant du gène UBBP4 reclassant ainsi ce dernier comme gène fonctionnel. Tous ces variants ont la caractéristique de présenter des mutations comparativement à la séquence de l'ubiquitine canonique. Plus précisément, l'un des variants, que nous avons nommé Ub<sup>KEKS</sup>, présente des différences fonctionnelles comparativement à l'ubiquitine en ce qui regarde la dégradation protéasomale. Par spectrométrie de masse, nous avons identifié les cibles de cette nouvelle modification post-traductionnelle qui présentent des cibles communes avec l'ubiquitine. Finalement, des cellules inactivées pour le gène UBBP4 présentent un retard de croissance avec la présence de nucléole significativement plus gros. La découverte de ces nouveaux variant, de leur cible et de la mise en évidence de phénotype dû à la délétion de ce gène ouvre la porte sur un nouvel aspect des modifications post-traductionnelles.

**Mots clés :** Ubiquitine, CRL4, DDB1, UBBP4, pseudogène, Ub<sup>KEK</sup>

## Table des matières

Résumé .....	v
Liste des figures .....	xi
Liste des tableaux .....	xiii
Liste des abréviations .....	xiv
<b>1. Introduction .....</b>	<b>1</b>
1.1 Découverte de l'ubiquitine .....	2
1.2 Évolution de l'ubiquitine .....	4
1.3 Gène de l'ubiquitine .....	5
1.4 Ubiquitine: Généralités .....	8
1.4.1 Mono-ubiquitylation .....	9
1.4.2 Poly-ubiquitylation .....	10
1.4.2.1 Met1 .....	11
1.4.2.2 Lysine 6 .....	12
1.4.2.3 Lysine 11 .....	12
1.4.2.4 Lysine 27 .....	13
1.4.2.5 Lysine 29 .....	13
1.4.2.6 Lysine 33 .....	14
1.4.2.7 Lysine 48 .....	14
1.4.2.8 Lysine 63 .....	15
1.4.3 Chaînes hétérogènes .....	15
1.4.4 Modification post-traductionnelle sur l'ubiquitine .....	18
1.4.4.1 Phosphorylation .....	19
1.4.4.2 Acétylation .....	19
1.4.4.3 SUMOylation .....	19
1.4.4.4 ADP-ribosylation .....	20
1.5 Ubiquitine-like .....	20
1.5.1 SUMO .....	21
1.5.2 NEDD8 .....	22
1.5.3 ATG8 .....	23
1.5.4 ATG12 .....	23

1.5.5 URM1.....	23
1.5.6 UFM1 .....	24
1.5.7 FAT10 .....	24
1.5.8 ISG15 .....	25
1.6 Reconnaissance de l'ubiquitine .....	25
1.7 Cascade enzymatique .....	28
1.7.1 La famille des enzymes E1 .....	29
1.7.2 La famille des enzymes E2 .....	30
1.7.3 La famille des enzymes E3 .....	32
1.8 Famille des RING.....	33
1.8.1 Complexe CRLs.....	34
1.8.1.1 Protéine RBX .....	35
1.8.1.2 Cullin.....	35
1.8.1.3 Protéines adaptatrices et protéines de reconnaissance .....	36
1.8.2 Régulation des complexes CRL .....	37
1.8.2.1 NEDDylation .....	37
1.8.2.2 CAND1 .....	37
1.9 Complexe CRL4 .....	38
1.9.1 Cul4A et Cul4B .....	38
1.9.2 DDB1 .....	39
1.9.3 DCAFs.....	40
1.10 Hypothèse et objectifs .....	41
1.10.1 Chapitre 1 : Étude des cibles du complexe DDB1-Cul4 par spectrométrie de masse via l'intermédiaire des protéines DCAFs.....	41
1.10.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène.....	42
<b>2. Matériel et méthode .....</b>	<b>43</b>
2.1 Général.....	43
2.1.1 Analyse de l'expression des ARNm à partir d'une base de données .....	43
2.1.2 Culture cellulaire, transfection et milieux SILAC .....	43
2.1.3 Extraits cellulaires et immunobuvardage.....	44
2.1.4 Immunofluorescence.....	45

2.1.5 Digestion des protéines.....	46
2.1.6 Spectrométrie de masse.....	47
2.1.7 Quantification et analyse bio-informatique.....	47
2.2 Spécifique au Chapitre 1 .....	48
2.2.1 Clonage et plasmide des protéines DCAFs .....	48
2.2.2 Co-immunoprécipitations pour DDB1-GFP.....	49
2.2.3 Purification par affinité à l'aide de BirA* .....	49
2.2.4 Pulse SILAC et digestion en solution .....	50
2.3 Spécifique au Chapitre 2 .....	51
2.3.1 Séquençage de UBBP4 pour l'ADN génomique et de l'ARNm.....	51
2.3.2 Identification des peptides provenant du gène de UBBP4 .....	52
2.3.3 Clonage et plasmides .....	52
2.3.4 Traitement au MG132 .....	52
2.3.5 Co-immunoprécipitations pour PCNA, LMNA et LMMB2 .....	53
2.3.6 CRISPR-Cas9 UBBP4 Knockout (KO).....	53
2.3.7 Essai de prolifération par Carboxyfluorescein succinimidyl ester (CFSE) .....	54
2.3.8 Quantification des nucléoles .....	54
2.3.9 Immunoprécipitation des protéines modifiées par HA-Ub et HA-Ub <sup>KEKS</sup> .....	54
2.3.10 Enrichissement des peptides modifiés avec GlyGly .....	55
<b>3. Résultats.....</b>	<b>56</b>
3.1 Chapitre 1 : Étude des cibles du complexe DDB1-Cul4 par spectrométrie de masse via l'intermédiaire des protéines DCAFs .....	56
3.1.1 La famille des protéines DCAFs.....	58
3.1.2 Profil des protéines DCAFs.....	59
3.1.2.1 Expression des ARNm des DCAFs.....	59
3.1.2.2 Identification des protéines DCAFs essentielles à la survie cellulaire .....	59
3.1.3 Identification des protéines qui interagissent avec la protéine DDB1 ....	61
3.1.4 Génération des lignées stables.....	63
3.1.5 Purification par affinité de l'interactome des DCAF .....	64
3.1.6 Interaction entre les DCAFs et les protéines du complexe CRL4 .....	66



3.1.7 Fonctions associées aux protéines DCAFs .....	68
3.1.8 Pulse SILAC.....	70
3.1.9 Exemple de DDB2.....	72
3.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène.....	76
3.2.1 Découverte d'UBBP4-B1 .....	77
3.2.2 Expression de l'ARNm du gène UBBP4 dans les cellules humaines.....	80
3.2.3 Traduction de l'ARNm d'UBBP4 .....	82
3.2.4 Modification du protéome par les variants de l'ubiquitine.....	84
3.2.5 Réponse d'UBBP4 à l'inhibition du protéasome.....	85
3.2.6 Fonction d'Ub <sup>KEKS</sup> .....	88
3.2.7 PCNA-UB <sup>KEKS</sup> .....	92
3.2.8 Délétion du gène d'Ub <sup>KEKS</sup> .....	94
3.2.9 Modification du signal di-Glycines dans les cellules Ub <sup>KEKS</sup> /- .....	96
3.2.10 Les lamines sont des cibles d'Ub <sup>KEKS</sup> .....	98
<b>4 Discussion.....</b>	<b>102</b>
4.1 Chapitre 1 : Étude des cibles du complexe DDB1-CulA via l'intermédiaire des protéines DCAFs.....	102
4.1.1 Profil des protéines DCAFs.....	104
4.1.2 Partenaires d'interaction.....	106
4.1.2.1 Immunoprécipitations de la protéine DDB1 .....	106
4.1.2.2 Confirmation de l'interaction des protéines DCAFs avec le complexe DDB1-Cul4.....	107
4.1.2.2 Identification des fonctions principales des protéines DCAFs. ....	111
4.1.3 Identification des cibles des protéines DCAFs. ....	113
4.1.4 DDB2 et TCERG1 .....	113
4.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène.....	117
4.2.1 Découverte des variants d'ubiquitine produit à partir de UBBP4 .....	117
4.2.2 Réponse des variants A2 et B1 à l'inhibition du protéasome.....	123
4.2.3 PCNA, FANCI et FANCD2 .....	124
4.2.4 Lamine et Ub <sup>KEKS</sup> .....	126

<b>5. Conclusion .....</b>	<b>129</b>
Chapitre 1 : Étude des cibles du complexe DDB1-CulA via l'intermédiaire des protéines DCAFs.....	129
Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène.....	131
<b>6. Bibliographie .....</b>	<b>134</b>

## Liste des figures

Figure 1 : Schéma des fonctions cellulaires de l'ubiquitine.....	2
Figure 2: Schéma des gènes de l'ubiquitine.....	6
Figure 3: Mono-Ubiquitinylation.....	10
Figure 4 : Séquence de l'ubiquitine.....	11
Figure 5: Type de chaînes de poly-ubiquitine .....	16
Figure 6 : Avantage des chaînes hétérogènes .....	18
Figure 7 : Cascade enzymatique.....	29
Figure 8 : Représentation des 3 grandes familles de E3 .....	33
Figure 9 : Composition des complexes CRLs.....	34
Figure 10 : Composition des différents complexes Cullin .....	36
Figure 11 : Profil d'expression et caractère essentiel des DCAFs .....	60
Figure 12 : Identification des protéines DCAFs qui interagissent avec DDB1 .....	63
Figure 13 : Preuve d'expression des protéines DCAFs-BirA* .....	64
Figure 14 : Identification des protéines qui interagissent de manière significative avec les protéines DCAFs. ....	65
Figure 15 : Heatmap des interactions entre les DCAFs et les composantes du complexe CRL4.....	67
Figure 16: Fonctions cellulaires associées aux complexes DDB1-Cul4-DCAFs .....	69
Figure 17 : Schéma de l'approche Pulse-SILAC.....	71
Figure 18 : Pulse SILAC des protéines DCAFs .....	72
Figure 19 : Confirmation des lignées cellulaires DDB2-BirA* .....	73
Figure 20 : Interactome de la protéine DDB2-BirA* .....	74
Figure 21 : Schéma de l'approche Pulse-SILAC.....	75
Figure 22 : Découverte d'une protéine alternative IP_710742 .....	77
Figure 23 : Le gène UBBP4 ne possède pas de codon-stop et n'est donc pas un pseudogène .....	79
Figure 24 : Expression des ARNm de la famille UBBP et d'UBB .....	81

Figure 25 : Preuve de l'expression de 4 différentes protéines à partir du gène UBBP4 dans deux cadres de lecture différents .....	83
Figure 26 : Expression des différentes protéines encodées dans le gène UBBP4 ..	85
Figure 27 : Réponse de l'ubiquitine, de UBBP4-A2 et B1 à un inhibiteur du protéasome le MG132 .....	86
Figure 28 : Immunofluorescence de l'ubiquitine et de UBBP4-B1 à la suite d'un traitement avec un inhibiteur du protéasome.....	87
Figure 29 : Provenance du nom Ub <sup>KEKS</sup> .....	88
Figure 30 : Approche expérimentale pour les expériences de spectrométrie de masse .....	89
Figure 31 : Protéines enrichies en conditions natives .....	90
Figure 32 : Protéines enrichies en condition dénaturante .....	91
Figure 33 : Modification post-traductionnelle de PCNA par HA-Ub ou de HA-Ub <sup>KEKS</sup> .....	93
Figure 34 : Modification post-traductionnelle des mutants de PCNA par HA-Ub ou par HA-Ub <sup>KEKS</sup> .....	94
Figure 35 : Stratégies utilisées pour les lignées UBBP4 <sup>-/-</sup> .....	95
Figure 36 : courbe de croissance des cellules UBBP4 <sup>-/-</sup> .....	96
Figure 37: schéma représentatif de l'approche utilisée pour identifier les cibles endogènes de Ub <sup>KEKS</sup> .....	97
Figure 38 : Résultats de l'analyse de spectrométrie de masse des protéines avec une diminution de la marque diglycines dans les cellules UBBP4 <sup>-/-</sup> .....	98
Figure 39 : Modification post-traductionnelle de LMNB2 et LMNA par HA-Ub ou par HA-Ub <sup>KEKS</sup> .....	99
Figure 40 : Localisation de LMNA dans les cellules UBBP4 <sup>-/-</sup> .....	100
Figure 41 : Quantification de l'aire nucléoles dans les cellules UBBP4 <sup>-/-</sup> .....	101
Figure 42 : Réseau d'interaction des protéines avec DDB1 selon leur degré de confiance .....	110
Figure 43: Fonctions du complexe DDB1-Cul4 .....	112

## Liste des tableaux

Tableau 1 : Famille des protéines ubiquitine-like .....	21
Tableau 2 : Les différentes classes de Ubiquitin binding domain .....	27
Tableau 3 : Anticorps primaires et secondaires utilisés.....	45
Tableau 4 : Amorces utilisées pour le séquençage de l'ARMm et du gène de UBBP4 .....	51
Tableau 5 :Tableau des protéines DCAFs .....	58

## Liste des abréviations

<b>ADN</b>	Acide désoxyribonucléique
<b>ADNc</b>	ADN complémentaire
<b>ARN</b>	Acide ribonucléique
<b>ARNm</b>	ARN messenger
<b>ATP</b>	Adenosine triphosphate
<b>BF</b>	Bayes Factor
<b>CFSE</b>	Carboxyfluorescein succinimidyl ester
<b>CRLs</b>	Cullin ring ligase
<b>CUL</b>	Cullin
<b>DAPI</b>	4.6-diamidino-2-phénylindole
<b>DCAF</b>	DDB1-Cullin4 associated factors
<b>DDB1</b>	DNA damage binding protein 1
<b>DMEM</b>	Dulbecco's modified eagle medium
<b>DMSO</b>	Diméthylsulfoxyde
<b>DTT</b>	Dithiothréitol
<b>DUBs</b>	Deubiquitinase
<b>FA</b>	Acide formique
<b>FACS</b>	Fluorescence-activated cell sorting
<b>FBS</b>	Fetal bovin serum
<b>FDR</b>	False Discovery Rate
<b>GFP</b>	Green Fluorescent Protein
<b>Gly</b>	Glycine
<b>HECT</b>	Homologous to the E6-AP Carboxyl Terminus
<b>IP</b>	Immunoprécipitation
<b>K</b>	Lysine
<b>kDA</b>	Kilo Dalton
<b>M</b>	Methionine
<b>MCM</b>	Minichromosome Maintenance
<b>ORF</b>	Open reading frame
<b>altORF</b>	alternative Open Reading Frame
<b>PBS</b>	Phosphate buffer saline
<b>PCNA</b>	Proliferating cell nuclear antigen
<b>PCR</b>	Polymerase Chain Reaction
<b>PTM</b>	Modification post-traductionnelle
<b>RBR</b>	Ring in Between Ring
<b>RING</b>	Really interesting new gene
<b>RT-PCR</b>	Reverse transcription polymerase chain reaction
<b>SDS</b>	Sodium dodecyl sulfate
<b>SILAC</b>	Stable isotope labeling by amino acids in cell culture

<b>SUMO</b>	Small Ubiquitin-like MOdifier
<b>TFA</b>	Acide trifluoroacétique
<b>TPM</b>	Transcript per million
<b>UV</b>	Ultraviolet
<b>UB</b>	Ubiquitin
<b>UBBP</b>	Ubiquitin pseudogene
<b>UBD</b>	Ubiquitin binding domain
<b>UBL</b>	Ubiquitin like

## **1. Introduction**

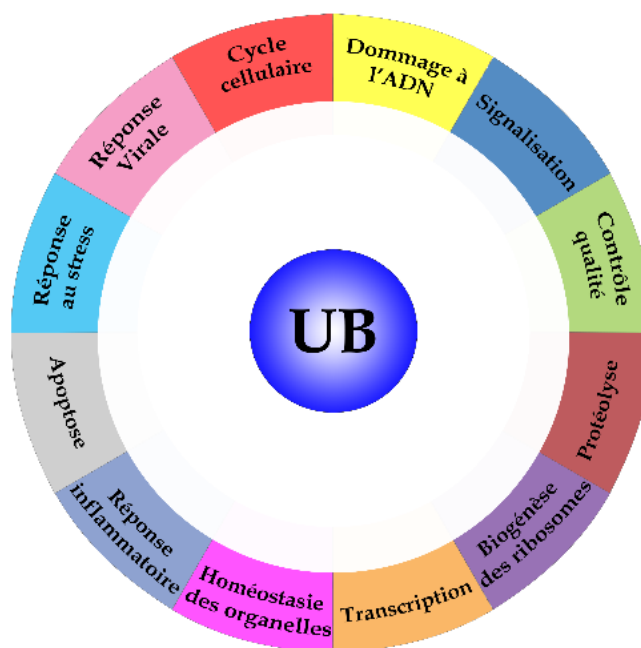
Les modifications post-traductionnelles (PTM) sont des groupements de différentes natures (tel que des composés chimiques ou protéiques) qui s'ajoutent sur la protéine avec l'aide d'enzymes tel que des kinases ou des méthyltransférases. Ces modifications influencent les caractéristiques biochimiques d'une protéine, ce qui a comme répercussion de moduler les fonctions cellulaires de celles-ci. La plupart de ces modifications ne sont pas statiques et vont être ajoutées et enlever de la protéine dépendamment du contexte cellulaire. Ainsi, grâce à ces modifications, la cellule va pouvoir utiliser les protéines déjà synthétisées pour produire une réponse rapide à divers stimuli.

Bien que la phosphorylation, la méthylation et l'acétylation aient été pendant de nombreuses années sous le phare des chercheurs, l'ubiquitine a gagné un intérêt considérable dans les dix dernières années. Ce gain de popularité peut être attribué à la grande versatilité fonctionnelle que peut avoir l'ubiquitine sur une protéine comparativement aux autres modifications. En effet, l'ubiquitine va s'attacher de manière covalente sur une lysine de la protéine cible mais puisque l'ubiquitine possède elle-même 7 lysines, elle peut être le propre sujet de sa conjugaison pour former des chaînes de poly-ubiquitine sur une protéine. Ainsi, les différentes lysines choisies, le type de chaînes et une panoplie d'autres facteurs vont conférer à cette modification une très grande amplitude dans ses fonctions (Figure 1). Il n'est donc pas étonnant d'observer que l'ubiquitine, dépendamment de sa conformation sur une protéine donnée, va pouvoir moduler les interactions, la localisation ou encore la stabilité de la protéine cible.

Aujourd'hui, l'ubiquitinylation est un domaine encore en expansion avec chaque année plusieurs milliers de publications qui tentent de découvrir toutes les régulations amenées par cette modification post-traductionnelle. Si initialement



l'ubiquitine était perçue comme un code qui conférait un rôle particulier à un type de chaînes données, cette idéologie a évolué pour désormais comparer l'ubiquitine à un langage complexe où la finalité du message est dépendant du contexte et de la nature de celui-ci.



**Figure 1 : Schéma des fonctions cellulaires de l'ubiquitine**

L'ubiquitine est une modification post-traductionnelle qui est connue pour son rôle dans la protéolyse. Aujourd'hui, l'implication de cette modification s'étend à presque toutes les fonctions cellulaires connues. En effet, que ce soit en condition physiologique ou pathologique, l'ubiquitine est au cœur d'un nombre incalculable de processus biologiques.

## 1.1 Découverte de l'ubiquitine

La protéolyse est le procédé enzymatique par lequel les protéines sont sujettes à une dégradation pour produire de plus petits peptides ou encore des acides aminés. Ce processus qui est hautement régulé est essentiel pour l'homéostasie cellulaire et on sait aujourd'hui que son altération entraîne rapidement la mort de la cellule. Pendant de très nombreuses années, il était accepté que la

plupart des protéines étaient relativement stables avec une longue demi-vie et que leur protéolyse dépendait uniquement de l'activité lysosomale, qui était à l'époque, un processus peu sophistiqué. Cette supposition a survécu jusqu'à la fin des années 1970 ou une série de découvertes est venue changer ce dogme. C'est Avram Hershko, Aaron Ciechanover et Irwin A. Rose qui ont décrit pour la première fois en 1978 une petite protéine appelée APF-1 qui était capable d'entraîner la dégradation de la protéine cible de manière ATP-dépendante (Ciechanover, Hod, et Hershko 1978). En 1980, ils ont émis l'hypothèse que APF-1 étaient en fait un marquage spécifique pour une protéase particulière nommée à l'époque APF-2 (Ciechanover et al. 1980). Ces deux découvertes successives venaient de mettre à jour pour la première fois le système d'ubiquitine (APF-1) protéasome (APF-2). La même année, ils ont décrit le processus de poly-ubiquitinylation et ont identifié l'activité enzymatique dé-ubiquitinase (DUB) qui est responsable de l'élimination de l'ubiquitine sur un substrat (Hershko et al. 1980). Finalement, durant les années qui suivirent, ils ont découvert les trois grandes classes d'enzymes responsables de la conjugaison de l'ubiquitine sur une protéine, soit les enzymes E1, E2 et E3 (A. L. Haas et al. 1982). Ainsi, en seulement quelques années, tous les principaux acteurs du système d'ubiquitine protéasome ont été identifiés par un nombre très restreint de chercheurs. Cette découverte a d'ailleurs valu aux trois chercheurs (Avram Hershko, Aaron Ciechanover et Irwin A. Rose) le prix Nobel de chimie en 2004 pour leurs travaux innovateurs dans la découverte de ce système protéolytique (K. D. Wilkinson 2004).

Néanmoins, ce système est resté catalogué comme un simple système de dégradation pour les protéines mal repliées, endommagées ou trop vieilles pendant plus d'une décennie. Il a fallu attendre le début des années 1990 pour que cette idée évolue avec entre autres, deux découvertes qui ont révélé l'importance de cette modification post-traductionnelle dans le contrôle de certains processus cellulaires. La première découverte marquante a été de constater que des protéines hautement régulées, telles que les cyclines, étaient des substrats du système d'ubiquitine-protéasome (Glutzer, Murray, et Kirschner 1991). En effet, puisque la quantité de

ces protéines est très importante lors du cycle cellulaire, la mise en évidence de leur dégradation par le protéasome venait de souligner l'importance de l'ubiquitine dans la régulation de ces protéines. La seconde découverte fût grâce au développement d'inhibiteur du protéasome, qui à la suite de leur utilisation, a permis de comprendre que le système d'ubiquitine protéasome était non seulement essentiel à la survie cellulaire, mais aussi impliqué dans presque tous les processus biologiques (Kisselev et Goldberg 2001).

Aujourd'hui près de 40 ans après la découverte de ce système, le domaine qui entoure l'ubiquitine reste encore aujourd'hui une source continue de découverte qui nous permet de mieux comprendre les processus biologiques et certaines pathologies telles que le cancer ou des maladies neurodégénératives dans lesquelles cette modification post-traductionnelle joue un rôle important.

## 1.2 Évolution de l'ubiquitine

Le système d'activation utilisé par l'ubiquitine a été pendant une longue période de temps associé exclusivement aux eucaryotes. En effet, tous les eucaryotes qu'ils soient unicellulaires, tels que les levures, ou qu'ils forment un système beaucoup plus complexe, tel que l'homme, possèdent un système d'ubiquitine-protéasome. Toutefois, des travaux réalisés par plusieurs équipes de recherches ont permis d'identifier des systèmes apparentés à l'ubiquitine (ubiquitine-like) chez les bactéries et les archaea (Lake et al. 2001). Chez la bactérie, les protéines MoaD et ThiS qui sont impliquées comme cofacteur à molybdène et dans la biosynthèse de la thiamine respectivement partagent une structure similaire à l'ubiquitine, possèdent deux glycines en C-terminal et sont activées par un mécanisme comparable (Leimkühler, Wuebbens, et Rajagopalan 2001) (Taylor et al. 1998). En ce qui concerne les archaea, la protéine SAMPs identifiée chez *Haloferax volcanii* joue un rôle dans le métabolisme du soufre et dans le stress oxydatif et elle possède aussi toutes les caractéristiques d'une signalisation ubiquitine-like (Humbard et al. 2010). Bien que les modifications chez les

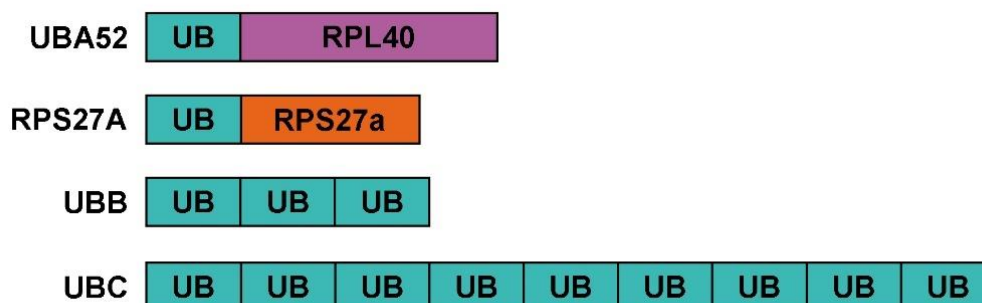
procaryotes utilisent les mêmes principes de conjugaison que l'ubiquitine, leur cascade d'activation reste tout de même beaucoup moins complexe que chez les eucaryotes, avec l'intervention d'une seule enzyme nécessaire pour leur activation. Plus récemment, lors du séquençage du génome de *Caldiarchaeum subterraneum*, un ensemble de gènes qui encode pour une cascade d'activation et de signalisation complète a été identifié avec un gène d'ubiquitine, un gène pour une E1, une E2, une E3 et même un gène avec une activité dé-ubiquitinase (Nunoura et al. 2011). Ainsi, l'ubiquitine n'est pas restreinte aux eucaryotes et serait probablement apparue assez rapidement dans l'évolution.

Un autre aspect très important de l'évolution de l'ubiquitine est la conservation étonnante de sa séquence protéique. Chez l'humain, l'ensemble des gènes de l'ubiquitine possèdent tous la même séquence en acide aminé qui est pratiquement la même chez presque tous les eucaryotes. En effet, l'alignement de séquence entre l'ubiquitine humaine et celle de la levure *Saccharomyces* ou encore à celle de la plante *Arabidopsis*, présente seulement trois acides aminés de différence entre chacune de ces séquences. L'une des théories quant à la conservation extraordinaire de cette protéine est que lors de l'apparition de l'ubiquitine chez les premiers organismes, il y aurait eu un gain de fonction important des mécanismes dépendant de cette modification, la rendant essentielle dans de nombreux processus cellulaires. Par conséquent, les caractéristiques biologiques et biochimiques de l'ubiquitine ont probablement été sélectionnées dans les premiers temps de l'évolution et sont restées les mêmes depuis.

### **1.3 Gène de l'ubiquitine**

Chez l'humain, l'ubiquitine est produite à partir de quatre gènes indépendants nommés UBA52, RPS27A, UBB et UBC qui peuvent se diviser en deux groupes en se basant sur leur structure. Les gènes d'ubiquitine monomérique, UBA52 et RPS27A produisent une seule molécule d'ubiquitine fusionnée à une protéine ribosomale et

les gènes de poly-ubiquitine, UBB et UBC produisent 3 et 9 molécules d'ubiquitine respectivement (Baker et Board 1991).



**Figure 2: Schéma des gènes de l'ubiquitine**

L'ubiquitine est produite dans la cellule humaine à partir de quatre gènes. Les deux premiers gènes produisent des protéines de fusion avec une protéine ribosomale essentielle alors que les deux autres gènes encodent pour une répétition de la séquence d'ubiquitine. Pour tous les gènes de cette protéine, l'activité DUBs est essentielle pour la libération des ubiquitines dans la cellule.

Pour les deux groupes de gènes, à la suite de la traduction de leur ARNm, les molécules d'ubiquitines sont soit fusionnées à une protéine ribosomale ou soit fusionnées entre elles. Ainsi, il est nécessaire de cliver spécifiquement après chacune des séquences d'ubiquitine pour permettre la libération de monomères fonctionnels. Ce processus est assuré par une famille d'enzyme appelée les dé-ubiquitinases (DUBs). L'identité de la DUB responsable du clivage des précurseurs n'a pas encore été parfaitement identifiée. Plusieurs études ont initialement attribué le rôle de clivage à UCHL1 et UCHL3 pour les gènes fusionnés aux protéines ribosomales (Larsen, Krantz, et Wilkinson 1998) et à USP5 pour les gènes de poly-ubiquitine (Falquet et al. 1995). Cependant, ces enzymes ne semblent pas être les seules à participer à ce processus puisque des souris qui ont les gènes UCHL1 et UCHL3 délétés demeurent viables (K.-Y. Ryu et al. 2010). Ceci démontre soit une compensation fonctionnelle par une autre DUB, soit que plusieurs membres de cette famille pourraient être responsables de ce mécanisme. Dans ce même ordre d'idée, une étude a démontré qu'en fait, un certain nombre de DUBs sont capables de

remplir cette fonction enzymatique mais que cela dépend non seulement du type de gène d'ubiquitine (ribosomal ou poly) mais aussi du type de tissu étudié (Grou et al. 2015). Un autre aspect de ce processus qui est encore matière à débat est le moment du clivage des précurseurs. La première hypothèse est que le clivage se fait de manière co-translationnelle alors que la seconde est que celui-ci serait plutôt un évènement post-translationnelle. Bien que les deux théories sont appuyées par diverses études, il semblerait que le clivage post-transcriptionnelle est fait de manière prédominante pour les gènes fusionnés aux protéines ribosomales et que les gènes de poly-ubiquitine seraient clivés par une combinaison des deux mécanismes (Grou et al. 2015).

Les deux gènes monomériques fusionnés à des protéines ribosomales sont essentiels et sont exprimés de manière constitutive ce qui assure à la cellule un niveau basal d'ubiquitine en tout temps (Baker et Board 1991). Les gènes poly-ubiquitine ont été pendant longtemps associés à des gènes de stress puisque leur expression est augmentée en réponse aux chocs thermiques, aux dommages à l'ADN et aux stress oxydatifs (Fornace et al. 1989). Cette régulation a amené à croire que les gènes de poly-ubiquitine contribuaient de manière moins importante au bassin d'ubiquitine en condition normale. Cependant, une étude a démontré que dans des lignées cellulaires humaines, 36% de l'ubiquitine total provenait de RPS27A, 24% provenaient du gène UBC, et que le 40% restant provenait de manière équivalente du gène UBA52 et de UBB. (Bianchi et al. 2015). Ainsi, contrairement à ce qui était accepté depuis de nombreuses années, les gènes de poly-ubiquitine produiraient près de la moitié de l'ubiquitine en condition basale. En contrepartie, les gènes monomériques semblent eux aussi pouvoir participer à la réponse aux stress cellulaires puisque leur ARNm peut être stabilisé à la suite de dommages à l'ADN (tel que les UV) laissant présager un rôle potentiel de ceux-ci dans ces mécanismes (Bianchi et al. 2015). Ainsi, contrairement aux premières théories émises quant aux gènes de l'ubiquitine, leur fonction et leur régulation semblent bien plus complexes que ce qui était initialement présumé.

Un autre aspect intéressant des gènes de l'ubiquitine et qui renforce leur caractère indépendant, est que bien que ceux-ci produisent la même séquence d'acides aminés, les rôles pour chacun ne semblent pas être redondants. Alors que la délétion des gènes RPS27A et UBA52 est impossible dû au caractère essentiel des protéines ribosomales produites par ceux-ci, plusieurs études se sont penchées sur les rôles des gènes UBB et UBC. Ces études ont démontré que la délétion du gène UBC chez la souris est létale à l'état embryonnaire dû très probablement au mauvais développement du foie (Kwon-Yul Ryu et al. 2007). De plus, ce phénotype n'est que très peu atténué par une réexpression ectopique de l'ubiquitine chez la souris, mettant en évidence la nécessité du gène et non seulement de la présence de la protéine. D'autre part, en absence du gène d'UBC, aucun des autres gènes de l'ubiquitine (mono ou poly-ubiquitine) ne semble augmenter son expression pour compenser la perte d'ubiquitine dans la cellule (Park, Yoon, et Ryu 2013). Les souris avec une délétion du gène UBB présentent pour leur part des problèmes métaboliques, des troubles du sommeil et une neuro-dégénération de l'hypothalamus (K.-Y. Ryu et al. 2010) (Kwon-Yul Ryu et al. 2008). Ainsi, bien que tous les gènes d'ubiquitine possèdent la même séquence protéique, plusieurs de leurs aspects, que ce soit leur clivage, leur expression ou encore leur fonction, semblent être propres à chacun. Ceci vient renforcer l'hypothèse que tous ces gènes pourraient contribuer de manière différente à l'homéostasie cellulaire.

## **1.4 Ubiquitine: Généralités**

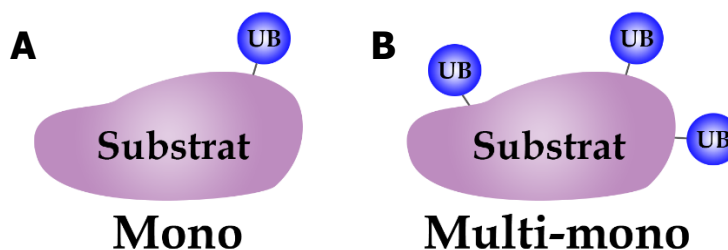
L'ubiquitine est une petite protéine très stable de 76 acides aminés avec un poids d'environ 8,6kDa. Celle-ci possède un domaine  $\beta$ -graps qui est formé de 2 hélices alpha et 5 feuillets bêta et d'une région plus flexible en C-terminal. La séquence peptidique de celle-ci se termine par deux glycines qui sont essentielles pour l'attachement de l'ubiquitine sur une protéine cible. L'ajout d'une ubiquitine se fait dans la grande majorité des cas sur le groupement  $\epsilon$ -amine d'une lysine interne au substrat pour former un lien thioester. Il existe cependant des exemples que

l'ubiquitine peu s'attacher sur un groupement  $\alpha$ -amine d'un résidu en N-terminal ou encore de manière plus rare sur les résidus cystéine, thréonine et sérine en formant d'autres types de liaison (McDowell et Philpott 2013). Ces ubiquitinylation non canoniques ont été démontrées cependant pour avoir les mêmes fonctions que celles canoniques telles qu'adresser les protéines au protéasome ou encore pour changer leur localisation cellulaire. Il existe deux théories qui pourraient expliquer l'existence de ce type de liaison. La première vient du fait que ce type de liaisons formées entre l'ubiquitine et ces acides aminés (C, S et T) sont de nature beaucoup plus faible que celles formées avec les lysines en plus de pouvoir se former plus rapidement. Ainsi dans des systèmes très dynamiques, la rapidité et la faiblesse de l'interaction pourraient être un avantage considérable. La deuxième hypothèse est que ce type de liaison serait utilisé par des protéines qui sont pauvres en lysine ou lorsque celle-ci ne sont pas accessibles pour la réaction d'ubiquitinylation. Ceci permettrait une régulation du substrat par l'ubiquitine pour la dégradation au protéasome (Ben-Saadon et al. 2004).

#### **1.4.1 Mono-ubiquitinylation**

Dans la version la plus simple de la réaction d'ubiquitinylation, un substrat est modifié par une ubiquitine sur une lysine pour amener ce qu'on appelle une mono-ubiquitinylation (figure 3A). Dans le même ordre d'idée un substrat peut être modifié sur plusieurs lysines pour produire une multi-mono-ubiquitinylation (figure 3B). Ces types d'ubiquitinylation vont surtout venir altérer les liens intra- ou intermoléculaires de la protéine afin de modifier les caractéristiques biochimiques de la protéine cible.





**Figure 3: Mono-Ubiquitylation**

L'ubiquitine peut s'attacher sur ses substrats de plusieurs manières. La plus simple est celle schématisée en A) où celle-ci forme une mono-ubiquitylation sur la protéine cible. Lorsqu'un substrat est la cible de plusieurs modifications de ce genre on parle plutôt de multi-mono-ubiquitine B)

### 1.4.2 Poly-ubiquitylation

L'une des particularités de l'ubiquitine est que celle-ci peut devenir le substrat de sa propre réaction enzymatique. En effet, puisque celle-ci possède 7 lysines dans sa séquence peptidique (Figure 4), elle peut elle-même se faire ubiquitinyler pour former ce qu'on appelle des chaînes de poly-ubiquitine. Les deux lysines les plus couramment utilisées sont K48 et K63 qui forment des chaînes de poly-ubiquitine typiques. Les autres lysines, soit K6, K11, K27, K29 et K33 forment des chaînes dites atypiques et sont beaucoup moins courantes au sein de la cellule (Akutsu, Dikic, et Bremm 2016). Toutefois, toutes les lysines de l'ubiquitine ont déjà été identifiées comme étant ubiquitinylées de manière endogène dans la cellule. De plus, il existe un 8<sup>e</sup> type de chaîne de poly-ubiquitine atypique qui est formé à partir de la méthionine en position 1 de l'ubiquitine. Lorsque les chaînes se construisent toujours sur la même lysine de l'ubiquitine, on parle alors de chaînes homogènes (Figure 5). Les chaînes homogènes possèdent toutes de nombreux rôles avec une redondance pour une même fonction cellulaire donnée. Par exemple, les chaînes de poly-K48 vont adresser les protéines au protéasome ainsi que les chaînes de poly-K11. Néanmoins, les chaînes vont avoir des rôles préférentiels dans leur fonction au sein de la cellule.

1 6 11 27 29 33 48 63  
 MQIFVKTTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRIRGG

#### Figure 4 : Séquence de l'ubiquitine

Séquence en acide aminé de l'ubiquitine. Les acides aminés en rouge sont ceux sur lesquelles les chaînes de poly-ubiquitine sont formées. Leur position dans la séquence est indiquée au-dessus puisque c'est de cette manière qu'on les nomme en parlant d'un type de lien spécifique.

##### 1.4.2.1 Met1

Les chaînes initiées sur la méthionine 1 sont certainement les liaisons atypiques qui sont le plus étudiées et pour qui la fonction a été largement décrite au cours des dernières années. Ce type d'attachement mène à une chaîne possédant une conformation linéaire. Cette forme de poly-ubiquitine a été découverte au milieu des années 2000 en même temps que l'enzyme E3 formée des protéines HOIP et HOIL-1, responsable de la synthèse de ces chaînes (Kirisako et al. 2006). Quelques années plus tard, un troisième membre s'est ajouté aux deux autres protéines formant la E3, la protéine SHARPIN, pour constituer le complexe LUBAC (Ikeda et al. 2011). En plus d'avoir un complexe enzymatique qui lui est propre et très bien caractérisé, son implication dans la réponse inflammatoire a été démontrée par de nombreuses équipes de recherche. Son rôle dans l'inflammation a été mis en évidence pour la première fois dans la signalisation engendrée par le récepteur du TNF (T. L. Haas et al. 2009). Par la suite, ce type de chaînes a été montré pour être impliqué dans la réponse médiée par les récepteurs à l'interleukine 1 (Emmerich et al. 2013a), par les récepteurs des cellules B (Gerlach et al. 2011) et par les senseurs intracellulaires de la présence bactérienne NOD2 (Damgaard et al. 2012). Pour l'ensemble de ces voies de signalisation, le recrutement de LUBAC entraîne la formation de chaînes Met1 sur une protéine cible servant de plateforme de recrutement pour d'autres effecteurs afin d'engendrer une cascade de signalisation menant à une réponse immunitaire adéquate.

#### **1.4.2.2 Lysine 6**

Si pendant longtemps les chaînes K6 n'avaient pas de rôle bien défini, plusieurs publications dans les dernières années sont venues mettre en évidence l'importance de ces chaînes dans 2 grands processus cellulaires. La première fonction de ce type de liaison a été son implication dans la réponse aux dommages à l'ADN. Une étude de spectrométrie de masse a démontré que suite à l'induction de dommage, il y avait un enrichissement des chaînes K6 (ainsi que des chaînes K33) dans la cellule (Elia et al. 2015). Cette étude venait faire écho à une autre qui avait démontré que la protéine BRCA1, une E3 ubiquitine ligase importante dans la réparation de l'ADN, pouvait former des chaînes poly-K6 lors de la réplication et de la réparation de l'ADN (Morris et Solomon 2004). L'autre fonction, qui est maintenant la plus connue pour ce type de liaison, est son rôle dans l'homéostasie mitochondriale. Plusieurs études ont démontré que la mitophagie médiée par la E3 Parkin, était régulée, entre autres, par des chaînes K6 (Durcan et al. 2014) (Ordureau et al. 2015). Finalement, un autre rôle pour les chaînes de poly-K6 a été dernièrement identifié ou une étude a démontré que la formation de chaînes K6 sur la protéine FADD (Fas-associated death domain) inhibait la formation du complexe DISC (death inducing signaling complex) et prévenait ainsi l'apoptose (Seo et al. 2018).

#### **1.4.2.3 Lysine 11**

Les chaînes de poly-K11 sont pour leur part impliquées, en parallèle avec les chaînes K48, dans la dégradation des protéines par le protéasome. Cependant, contrairement aux chaînes K48 qui sont impliquées à plus grande échelle dans la protéolyse, les liaisons K11 semblent être impliquées de manière plus spécifique dans la régulation du cycle cellulaire. En effet, le complexe d'ubiquitine ligase APC/C qui est essentiel à la progression de la phase mitotique du cycle cellulaire, utilise des chaînes K11 pour envoyer ces substrats à la dégradation au protéasome (L. Jin et al. 2008a). Cependant, des études ont démontré que les chaînes branchées K11/K48 qui seraient responsable de la dégradation de ces protéines au détriment

des chaînes homogènes K11 (Meyer et Rape 2014). Un autre exemple du rôle des chaînes mixte avec la lysine 11 est le rôle décrit des branchements mixtes K11/K63 dans l'endocytose du complexe majeur d'histocompatibilité de classe 1 ou les deux lysines sont importantes pour l'internalisation du complexe (Boname et al. 2010).

#### **1.4.2.4 Lysine 27**

De toutes les chaînes de poly-ubiquitine, c'est certainement celles formées sur la lysine 27 qui restent encore les moins bien caractérisées. Seulement quelques études ont rapporté la formation de ce type de chaîne *in vivo* et son implication dans des fonctions biologiques. Cependant, toutes les publications semblent faire état d'un rôle dans les interactions protéine-protéine pour les chaînes K27. En effet, c'est en 2014 que pour la première fois des chercheurs ont démontré que la poly-ubiquitinylation K27 d'un des récepteurs de l'autophagie sélective (OPTN) est nécessaire pour l'interaction avec d'autres protéines essentielles au processus autophagique (Z. Liu et al. 2014). De plus, une seconde publication a rapporté la formation de chaînes K27 suite à l'induction de cassures double-brin à l'ADN sur l'histone H2A/H2A.x qui serait nécessaire pour le recrutement des facteurs de réparation aux foyers de dommages (Gatti et al. 2015). Mettant encore une fois l'emphasis sur le fait que la formation de chaîne K27 permet l'association entre le substrat ubiquitinylé et ses partenaires d'interaction. Finalement, une dernière protéine a été démontrée pour être le sujet d'une poly-ubiquitinylation K27, et ce, dans les voies de signalisation MAP Kinase. En effet, la protéine BRAF pourrait être poly-ubiquitinylée sur la lysine 27 afin d'empêcher sa liaison avec un inhibiteur de son activité kinase laissant la voie de signalisation active (Yin et al. 2019).

#### **1.4.2.5 Lysine 29**

En condition basale, de tous les liens atypiques ce sont les chaînes K29 qui sont les plus abondantes (Dammer et al. 2011). Suite à l'inhibition du protéasome, la quantité de ces chaînes augmente dans la cellule et a donc initialement été rapportée pour jouer un rôle dans ce processus cellulaire (E. S. Johnson et al. 1995).

Dans les années suivantes, ces chaînes ont été impliquées dans le contrôle des voies de signalisation en même temps que des chaînes K33 (Al-Hakim et al. 2008). Plus récemment, des études ont démontré que des liens K29 et K27 peuvent se former sur des protéines et mener à leur agrégation dans la cellule (Nucifora Jr et al. 2016). Hormis ces quelques études qui ont démontré l'implication de ces chaînes dans différents processus cellulaires, très peu de choses sont connues pour ce type de liaison.

#### **1.4.2.6 Lysine 33**

Comme mentionné précédemment l'un des premiers rôles identifiés pour les chaînes K33 a été dans le contrôle des voies de signalisation associées à la protéine *AMPK-related kinase* (ARK) au même titre que les chaînes de poly-K29 (Al-Hakim et al. 2008). Un autre rôle identifié pour ces chaînes a été aussi mentionné dans la description des chaînes K6 où dans la même étude les auteurs avaient démontré une augmentation des chaînes K33 et K6 en réponse aux dommages à l'ADN (Elia et al. 2015). L'un des rôles plus spécifiques à ce type de chaînes est son implication dans le transport vésiculaire de l'appareil de Golgi au réticulum endoplasmique via la modulation des interactions protéine-protéine (W.-C. Yuan et al. 2014).

#### **1.4.2.7 Lysine 48**

De toutes les chaînes de poly-ubiquitine formées dans la cellule, les chaînes K48 sont les plus abondantes et les mieux caractérisées puisque ces liaisons sont le signal majoritaire pour l'envoi des protéines à la dégradation par le protéasome. Cependant, il existe une certaine complexité dans les signaux de dégradation que peuvent engendrer ce type de chaînes. En effet, si pendant plusieurs années le dogme d'une seule chaîne composée de 3 ubiquitines était le signal minimal requis pour l'envoi au protéasome, des études récentes ont démontré que deux chaînes de diubiquitine K48 sur un substrat étaient un meilleur signal de dégradation (Lu et al. 2015). Ainsi, la longueur des chaînes ne serait pas nécessairement le point le plus important en ce qui concerne l'envoi au protéasome. Toutefois, même si la poly-

ubiquitylation K48 est dans presque tous les cas associés à la dégradation, il existe des évidences qui démontrent que ces liaisons pourraient posséder d'autres fonctions biologiques. Cependant, l'un des seuls exemples rapportés est pour le facteur de transcription Met4 chez la levure, où la formation de chaîne poly-K48 protégerait la protéine de la dégradation (Flick et al. 2006).

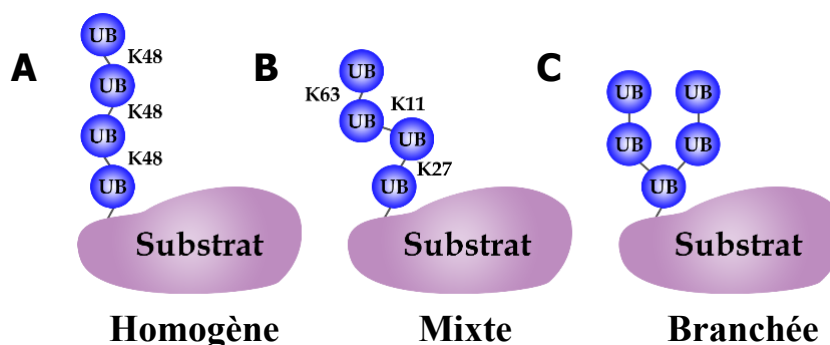
#### **1.4.2.8 Lysine 63**

Les chaînes de poly-K63 sont, avec les chaînes de poly-K48, les liaisons les plus étudiées et les mieux caractérisées de l'ubiquitine. Depuis de nombreuses années des études ont démontré le rôle de ces chaînes dans de nombreuses fonctions telle que la réparation des dommages à l'ADN, la réponse inflammatoire, la signalisation (plus particulièrement par l'activation de kinase) et dans bien d'autres fonctions (Husnjak et Dikic 2012). Dans la majorité des cas, ces chaînes servent de plateforme de reconnaissance pour les interactions protéines-protéines. En fait, ces chaînes sont bien connues pour avoir des fonctions complètement indépendantes du protéasome puisqu'elles ne présentent aucune accumulation suite à son inhibition (Xu et al. 2009) et qu'elles ne s'associent pas avec celui-ci (Jacobson et al. 2009). Cependant, il a été démontré que des chaînes K63 pouvaient être l'étape initiale pour la formation de chaînes hétérogènes K48/K63 afin d'adresser certains substrats spécifiques à la dégradation au protéasome (Ohtake et al. 2018). Ainsi, ces chaînes pourraient aussi servir de signal de dégradation lorsqu'elles sont en complexe avec d'autres types de chaînes.

#### **1.4.3 Chaînes hétérogènes**

Comme décrit précédemment, les chaînes de poly-ubiquitine peuvent être homogènes. Cependant, les chaînes hétérogènes qui se retrouvent au sein de la cellule peuvent se regrouper en deux groupes, soit les chaînes mixtes ou encore les chaînes branchées (Figure 5). Les chaînes mixtes impliquent que l'ubiquitine est modifiée sur une seule de ses lysines mais avec différentes liaisons alors que les

chaînes branchées sont lorsqu'une ubiquitine est modifiée sur 2 de ses lysines ou plus.



**Figure 5: Type de chaînes de poly-ubiquitine**

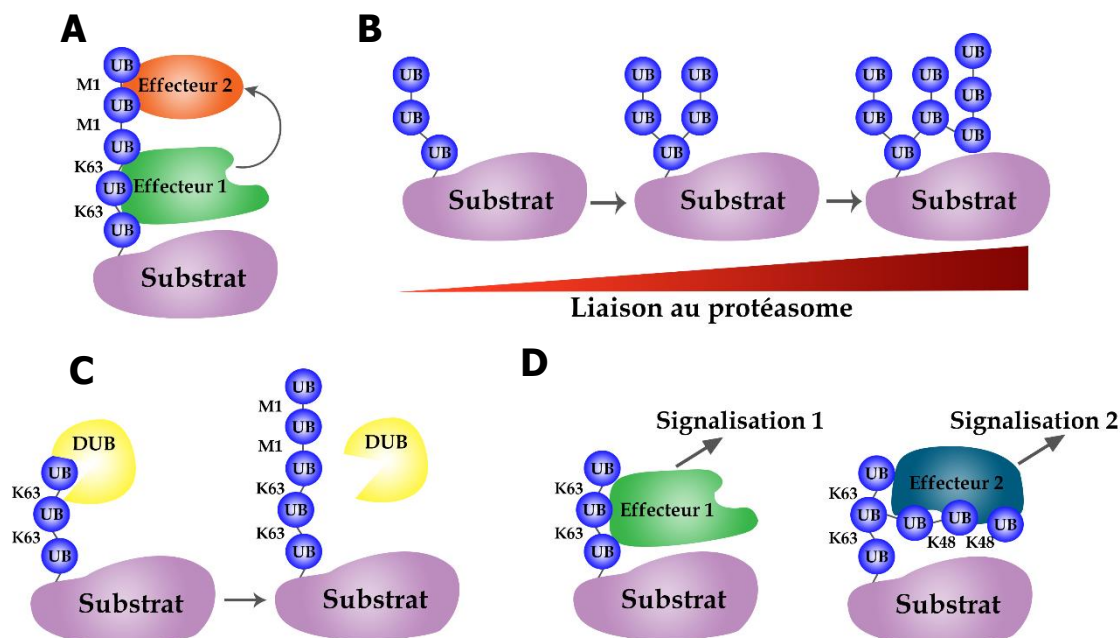
L'ubiquitine peut former une multitude de chaînes de poly-ubiquitine qui vont avoir toute une panoplie de fonctions cellulaires. En A, on retrouve les chaînes de poly-ubiquitine homogènes, ce sont les chaînes les plus abondantes dans la cellule. En B, on retrouve les chaînes hétérogènes mixtes qui sont formées à la suite de l'utilisation de différentes lysines lors de la synthèse de la chaîne. Finalement en C, une représentation des chaînes branchées qui vont se définir par l'utilisation de plusieurs lysines sur une ubiquitine pour former des chaînes à partir de celle-ci tel qu'illustrée.

Il existe plusieurs avantages biologiques à l'utilisation de chaînes mixtes et branchées pour la cellule comparativement aux simples chaînes homogènes. En effet, l'un des avantages des chaînes mixtes par exemple est le recrutement de plusieurs effecteurs via la même chaîne (Figure 6A). Un bon exemple de ce procédé sont les chaînes K63/M1 formées sur le récepteur de l'interleukine-1. Lors de l'activation de ce récepteur, des chaînes de poly-K63 sont formées qui permettent le recrutement de la kinase TAK1. Par la suite, la chaîne d'ubiquitine formée se poursuit avec une liaison sur la M1 (formation de chaînes mixte) ce qui permet le recrutement du complexe IKK qui est activé par la kinase TAK1 pour engendrer une réponse inflammatoire (Emmerich et al. 2013b). L'utilisation de chaînes mixtes dans ce cas précis permet l'accélération de l'activation du complexe IKK par TAK1 par leur recrutement au même endroit via les deux types de chaînes d'ubiquitine et donc, une réponse cellulaire plus rapide à un stimulus inflammatoire. De plus, cette stratégie pourrait avoir un autre avantage où les chaînes M1 viennent protéger les

liens K63 de la dégradation par une DUB pour permettre une signalisation plus longue (Figure 6C) (Wertz et al. 2015). L'un des exemples de chaîne branchée le mieux caractérisé est celui mentionné précédemment pour le complexe APC/C qui forme des chaînes K11/K48 pour l'envoi des protéines qui régulent le cycle cellulaire à la dégradation au protéasome (Meyer et Rape 2014). En fait, une étude a démontré que les chaînes formées avec seulement des liens sur la lysine 11 seraient un mauvais substrat pour le protéasome mais que l'ajout de chaînes branchées K48 augmenterait considérablement la dégradation des protéines du cycle cellulaire (Figure 6B) (Grice et al. 2015). Inversement, l'inhibition de la formation de chaînes K11 entraîne un délai considérable dans la dégradation des protéines responsable de la sortie de la mitose amenant un retard dans la progression du cycle cellulaire (Min et al. 2015). Ainsi, la formation de chaînes branchées K11/K48 semble nécessaire pour la régulation du temps de vie de ces protéines. L'une des théories qui pourraient expliquer pourquoi ces chaînes sont un puissant signal de dégradation est que la concentration de chaînes d'ubiquitine capable de lier le protéasome est considérablement augmentée dans ce contexte ce qui permet un meilleur adressage de la protéine. Finalement, un autre avantage de l'utilisation des liaisons hétérogènes serait qu'elles peuvent signaler de manière indépendante l'une de l'autre (Nakasone et al. 2013). Dans le cas de la formation d'une chaîne branchée, la première chaîne formée peut recruter un premier effecteur afin d'engendrer une première cascade de signalisation. Dans un second temps, la formation d'une chaîne branchée sur la première permettrait de recruter un deuxième effecteur afin de produire une autre signalisation indépendante de la première (Figure 6D).

Somme toute, en considérant les lysines utilisées pour la formation des chaînes en plus de la composante homogène, mixte ou branchée, la diversité des signaux que peut envoyer l'ubiquitine sur une même protéine est pratiquement infinie. Ainsi, il est important de prendre en considération tous les éléments lors de la formation des chaînes sur une protéine afin de bien cerner la fonctionnalité de celle-ci sur son substrat.





**Figure 6 : Avantage des chaînes hétérogènes**

A) Les chaînes hétérogènes mixtes peuvent être utilisées pour recruter deux effecteurs différents sur la même protéine en utilisant la même chaîne ce qui permet d'être plus rapide et de construire une signalisation cellulaire ordonnée. B) Il a été démontré que l'utilisation des chaînes branchées permettait d'adresser certains substrats de manière plus efficace à la dégradation au protéasome entre autres, par l'augmentation du nombre de chaînes capables de lier le protéasome. C) Les chaînes mixtes pourraient aussi servir de protection contre l'attaque des DUBs sur des chaînes afin d'augmenter le temps de survie de celle-ci et donc de prolonger leur effet. D) Dans ce cas de figure, les deux effecteurs recrutés via les deux types de chaînes ne participent pas à la même voie de signalisation cellulaire mais permettent plutôt diverses réponses avec la même chaîne.

#### 1.4.4 Modification post-traductionnelle sur l'ubiquitine

Bien que l'ubiquitine soit une modification post-traductionnelle, celle-ci reste tout de même une protéine composée d'acides aminés qui peuvent être eux aussi sujets à des modifications. En effet, plusieurs modifications telles que la phosphorylation, l'acétylation, la SUMOylation ainsi que l'ADP-ribosylation ont déjà été détectées sur l'ubiquitine. La découverte que l'ubiquitine peut elle aussi être sujette à d'autres PTM est venu rajouter encore une fois une couche de complexité à la compréhension de la fonction de celle-ci sur le devenir des protéines cibles.

Cependant hormis quelques cas qui ont été plus étudiés (tel que pour la phosphorylation), le rôle de ces modifications post-traductionnelles restent encore à élucider.

#### **1.4.4.1 Phosphorylation**

Parmi celles qui ont été étudiées, la phosphorylation de la sérine 65 est connue pour avoir un rôle physiologique dans la mitophagie médiée par PINK1 (Koyano et al. 2014). Outre cet acide aminé, des études de spectrométrie de masse ont déterminé que l'ubiquitine peut être phosphorylée sur 8 des 11 sites potentiels (serines, thréonines et tyrosines) que contient sa séquence (Herhaus et Dikic 2015). Toutefois, aucune donnée n'est disponible quant aux kinases responsables de ces modifications ou des impacts sur les propriétés biochimiques de l'ubiquitine.

#### **1.4.4.2 Acétylation**

L'acétylation des lysines K6 et K48 a été détectée dans la cellule *in vivo* empêchant par le fait même la formation de chaînes sur ces deux résidus (Ohtake et al. 2015). Étrangement, celles-ci vont aussi interférer avec la formation de chaînes K11 et K63. Ce mécanisme pourrait notamment servir à la cellule à la régulation de la formation des chaînes de poly-ubiquitine sur un substrat.

#### **1.4.4.3 SUMOylation**

La SUMOylation de l'ubiquitine a aussi été détectée par spectrométrie de masse dans différentes études, et ce, sur plusieurs lysines (K6, K11, K27, K48 et K63) (Galisson et al. 2011; Lamoliatte et al. 2013). Même si la SUMOylation de l'ubiquitine semble augmenter suite à l'inhibition du protéasome et suite à un choc thermique, aucun rôle fonctionnel n'a encore été trouvé pour ce type de modification (Hendriks et al. 2014).

#### 1.4.4.4 ADP-ribosylation

Plus récemment, l'ajout d'ADP-ribose sur l'ubiquitine a été démontré par une équipe de chercheurs en 2017 (C.-S. Yang et al. 2017). Dans cette étude, ils ont démontré qu'un complexe d'ubiquitine ligase formé par Parp9/Dtx3L était capable de modifier l'ubiquitine par l'ajout d'un ADP-ribose sur sa dernière glycine en position 76 ce qui empêchait sa liaison à un substrat et donc bloquerait l'activité E3-ligase du complexe. Les auteurs ont démontré que ce mécanisme serait un contrôle négatif dans l'activité du complexe qui est impliqué dans la réparation des dommages à l'ADN pour que celui-ci soit efficace seulement en présence de bris double-brin. Ainsi, l'ajout d'ADP-ribose pourrait être un contrôle négatif pour certains complexes de E3-ubiquitine-ligase.

### 1.5 Ubiquitine-like

Bien que l'ubiquitine soit la première modification post-traductionnelle de ce genre à avoir été découverte, il existe un certain nombre de protéines qui s'apparente à l'ubiquitine du fait que ce sont elles aussi de petites protéines capables de s'attacher de manière covalente à leur substrat en utilisant une cascade enzymatique. Ces protéines sont structurellement et d'un point de vue évolutif, très semblables à l'ubiquitine, elles sont toutes regroupées dans la même famille de protéine appelée *ubiquitin-like* (Ubl). Ces protéines possèdent que très peu d'homologie de séquence avec l'ubiquitine (seulement 60% pour NEDD8 qui est celle qui possède le plus d'analogie avec l'ubiquitine) (Cajee, Hull, et Ntwasa 2012). C'est au niveau de leur structure tridimensionnelle que celles-ci se ressemblent avec une conformation centrée autour d'un domaine  $\beta$ -graps retrouvé chez tous les Ubl ainsi que chez l'ubiquitine. Chez l'homme, on retrouve 8 grandes familles de Ubl; SUMO, NEDD8, ATG8, ATG12, URM1, UFM1, FAT10 et ISG15. Si certaines de ces modifications se retrouvent chez tous les eucaryotes, certaines telles que FAT10 et ISG15, se retrouvent seulement chez les mammifères supérieurs dû entre autres à

leur rôle dans le système immunitaire (Farrell, Broeze, et Lengyel 1979) (Y.-C. Liu et al. 1999).

**Tableau 1 : Famille des protéines ubiquitine-like**

<b>Famille</b>	<b>Membres</b>	<b>Homologie avec l'ubiquitine</b>	<b>Référence</b>
<b>SUMO</b>	SUMO1, SUMO2, SUMO3, SUMO4	11-18%	(Pichler et al. 2017)
<b>NEDD8</b>	NEDD8	60%	(Enchev, Schulman, et Peter 2015)
<b>ATG8</b>	LC3A, LC3B, LC3B2, LC3C, GABARAP, GABARAPL1, GATE-16	10%	(Kriegenburg, Ungermann, et Reggiori 2018)
<b>ATG12</b>	ATG12	10%	(Murrow et Debnath 2015)
<b>URM1</b>	URM1	13%	(Cappadocia et Lima 2018)
<b>UFM1</b>	UFM1	14%	(Wei et Xu 2016)
<b>FAT10</b>	FAT10	29%	(Aichem et Groettrup 2016)
<b>ISG15</b>	ISG15	30%	(Dzimianski et al. 2019)

Toutes ces modifications possèdent leurs propres machineries enzymatiques composées de trois étapes qui leur permettent, tout comme l'ubiquitine, de s'attacher à leur substrat afin de réguler diverses fonctions cellulaires. Leur machinerie reste cependant moins complexe que l'ubiquitine en ce qui a trait au nombre des différentes enzymes retrouvées dans chacune des classes. Cependant, des exceptions existent qui démontrent que certaines de ces modifications peuvent dans un contexte particulier utiliser les enzymes qui sont associés majoritairement à l'ubiquitine.

### **1.5.1 SUMO**

De toutes les Ubl, ce sont certainement les protéines de la famille SUMO (pour *small-ubiquitin-related modifier*) composée de quatre membres (SUMO1, SUMO2,

SUMO3 et SUMO4) dont les fonctions ont été le plus étudiées. Les protéines SUMO2 et 3 partagent 97% d'homologie de séquence entre elles mais seulement 50% avec SUMO1. Cette faible ressemblance n'empêche toutefois pas les protéines de la famille SUMO à partager la même machinerie enzymatique (Tatham et al. 2001). Tout comme l'ubiquitine, les protéines SUMO peuvent former des chaînes composées de plusieurs molécules de SUMO à partir d'une lysine. Cependant, contrairement à l'ubiquitine qui peut former des chaînes sur toutes les lysines présentent dans sa séquence, SUMO2/3 ont une préférence pour la formation de chaînes sur la lysine 11. Cette famille de protéine a initialement été caractérisée pour moduler la localisation cellulaire de la protéine RanGAP1 lorsque celle-ci est SUMOylée (Matunis, Coutavas, et Blobel 1996). Depuis, elles ont été impliquées dans une multitude de fonctions cellulaires entre autres dans le transport et l'organisation nucléaires, la transcription, le remodelage de la chromatine, la réparation des dommages à l'ADN ainsi que la biogenèse des ribosomes (K. A. Wilkinson et Henley 2010).

### **1.5.2 NEDD8**

La NEDDylation a été initialement découverte pour moduler l'activité d'une classe de E3 qui est spécifique à l'ubiquitine, en modifiant tous les membres de la famille des Cullin afin de permettre leur activation (Liakopoulos et al. 1998). Par conséquent, les modifications apportées par la protéine NEDD8 semblent interreliées au système d'ubiquitine. Dans le même ordre d'idée, le récepteur à l'EGF (EGFR) est modifié par l'ubiquitine ainsi que par NEDD8 via la même E3, ce qui a pour conséquence d'accélérer sa dégradation (Oved et al. 2006). Cette étude a non seulement encore une fois prouvé la coopération fonctionnelle entre les deux systèmes, mais a aussi démontré que les deux modifications pouvaient utiliser les mêmes enzymes E3. En contrepartie, plusieurs protéines qui ont été identifiées comme étant NEDDylées sont plus stables lorsque modifiées; telle que la protéine HuR, des protéines ribosomales, la protéine LKB1 et AKT (Embade et al. 2012)

(Xirodimas et al. 2008) (Barbier-Torres et al. 2015). Ainsi, NEDD8 peut agir de manière indépendante de l'ubiquitine et possède ses propres fonctions cellulaires.

### **1.5.3 ATG8**

On dénombre chez l'homme 7 protéines qui font partie de la famille de ATG8 qui selon leur homologie sont séparées en deux, les LC3 (*mitrotubule-associated protein light chain 3*) et les GABARAP (*γ-aminobutyric acid receptor-associated protein*). Ces protéines diffèrent des autres Ubl puisque celles-ci ne ciblent pas les protéines mais plutôt se font modifier par des lipides et plus particulièrement des phosphatidyléthanolamines (PE). Ces modifications ont été largement étudiées dans le contexte de l'autophagie où elles interviennent dans tous les stades de la formation des autophagosomes (Kirisako et al. 1999).

### **1.5.4 ATG12**

ATG12, aussi impliqué dans la formation des autophagosome, contient uniquement une glycine en C-terminal qui permet son attachement à ATG5, un des deux seuls substrats connus à ce jour pour cette modification (Mizushima et al. 1998). La conjugaison d'ATG12 sur ATG5 permet leur interaction avec la protéine ATG16L1 afin de former un complexe de E3 nécessaire pour la conjugaison des Ubl de la famille ATG8 (Hanada et al. 2007). La seule autre fonction d'ATG12 qui a été rapportée est son rôle dans l'homéostasie mitochondriale et l'apoptose via sa conjugaison sur ATG3 qui est par le fait même sa E2 (Rubinstein et al. 2011) (Radoshevich et al. 2010).

### **1.5.5 URM1**

URM1 pour *ubiquitin-related modifier-1* fait partie de la famille des Ubl pour sa ressemblance structurale avec l'ubiquitine ainsi que par son besoin d'être activé par une enzyme E1. C'est aussi l'Ubl qui a été la plus conservée lors de l'évolution avec une très grande homologie avec les protéines MoaD et ThiS retrouvées chez les procaryotes (Furukawa et al. 2000). Celle-ci est principalement connue pour son

rôle comme donneur de soufre pour certaine réaction à la suite de son activation particulière par sa E1 tel que dans la thiomodification des ARNt (Schlieker et al. 2008). URM1 peut aussi modifier des protéines sur des lysines. Cependant, sa fonction comme Ubl plus classique est très peu étudiée et encore mal comprise. Des études à large spectre ont cependant identifié plusieurs cibles potentielles entre autres dans la réponse au stress oxydatif (Van der Veen et al. 2011).

### **1.5.6 UFM1**

La protéine UFM1 (*Ubiquitin fold modifier 1*) a été découverte en 2004 en même temps que les enzymes responsables de sa conjugaison (Komatsu et al. 2004). Les premières vraies évidences de la fonction de cette modification ont cependant été découvertes récemment. Si des études avaient été capables de lier UFM1 au stress du réticulum endoplasmique auparavant (Y. Zhang et al. 2012), c'est seulement dernièrement que le lien entre cette modification et le réticulum a été élucidé par la mise en évidence que la protéine ribosomale RPL26 doit être modifiée par UFM1 pour le bon fonctionnement de la traduction couplée au réticulum endoplasmique (Walczak et al. 2019).

### **1.5.7 FAT10**

L'Ubl FAT10 est présent uniquement chez les mammifères supérieurs et s'exprime dans le cytoplasme des cellules dendritiques matures et dans les cellules B de manières constitutives mais peut s'exprimer dans d'autres types cellulaires suite à différents stimuli du système immunitaire (Raasi et al. 1999). Tout comme l'ubiquitine, cette modification a été démontrée plusieurs fois pour mener ses substrats à la dégradation au protéasome (Hipp et al. 2005). D'autres études ont cependant démontré que cette modification pouvait avoir d'autres fonctions telles que la stabilisation de ses substrats (Xiuxia Liu et al. 2016) et pourrait jouer un rôle important dans la carcinogenèse (Aichem et Groettrup 2016).

### 1.5.8 ISG15

La première protéine Ubl à avoir été identifiée est l'ISG15 (*interferon-stimulated gene 15*) à la fin des années 1980 donc peu de temps après l'identification de l'ubiquitine (Farrell, Broeze, et Lengyel 1979). Tout comme FAT10, cette modification est absente des organismes inférieurs tels que la levure, les nématodes ou encore chez les drosophiles, ce qui indique que celle-ci est exprimée seulement chez les mammifères qui présentent un système immunitaire plus développé. L'expression de ISG15 est augmentée suite à la stimulation par l'interféron de type I (Loeb et Haas 1992), par les infections virales (W. Yuan et Krug 2001) et par les lipopolysaccharides (Malakhova et al. 2002). Seule, la protéine est capable d'être sécrétée par la cellule pour jouer un rôle ressemblant à celui des cytokines (D'Cunha et al. 1996). ISG15 peut aussi se conjuguer à différents substrats pour modifier leur interaction protéine-protéine ou encore empêcher leur dégradation par le protéasome. Cependant, le rôle exact de cette modification n'est pas encore très bien compris et semble dépendre du type d'infection ainsi que de l'espèce étudiée.

## 1.6 Reconnaissance de l'ubiquitine

L'ubiquitine est capable d'envoyer une multitude de signaux qui peuvent être interprétés par la cellule via des protéines qui possèdent un domaine capable de lire ces messages qu'on appelle les *Ubiquitin binding domain* (UBD). Ceux-ci vont reconnaître plusieurs des régions hydrophobes situées sur l'ubiquitine qui sont localisées à la surface de la protéine. La première région hydrophobes à avoir été caractérisée est celle qui se construit autour de l'Ile44, la Leu8 et de la Val70 et qui est couramment nommée Ile44 (Beal et al. 1996). Cette région est celle qui est reconnue par le protéasome et par une grande majorité des protéines possédant un UBD. Une seconde surface importante pour la reconnaissance de l'ubiquitine est celle composée de l'Ile36, la Leu71 et la Leu73 et qui se situe dans la région C-terminal de l'ubiquitine (Reyes-Turcu et al. 2006). Finalement, la dernière surface



hydrophobe est celle centrée autour de la Phe4 et qui est composée des résidus Gln2, Phe4, et Thr12 (Sloper-Mould et al. 2001). En plus des domaines hydrophobes, certains domaines de l'ubiquitine, tel que le TEK-box (Thr12, Thr14, Glu34, Lys6 et Lys11), ont été identifiés pour être nécessaires dans certaines fonctions telle que l'élongation de chaînes K11 (L. Jin et al. 2008b). Ainsi, dépendamment du type de liaison formé par l'ubiquitine, les surfaces de reconnaissance sont agencées de manière différente dans l'espace pour la spécificité du signal.

Comme mentionné, les modifications par mono-ubiquitine aussi bien que les chaînes de poly-ubiquitine sont reconnues par des protéines qui contiennent au moins un domaine UBD. Plusieurs UBD ont été découverts à ce jour et même si aucun consensus universel ne peut les définir, ils peuvent néanmoins se regrouper en différentes classes selon leur élément de liaison à l'ubiquitine composé soit d'hélice- $\alpha$ , doigt de zinc, ou domaine d'homologie avec Pleckstrin. Une autre classe regroupe des UBD qui partagent des similarités structurales avec les enzymes E2 mais qui n'ont pas d'activité catalytique (relatif à la conjugaison de l'Ub). Une dernière classe regroupe des UBD qui présentent d'autres types de domaines de reconnaissance et qui sont aujourd'hui encore les seuls à présenter ces éléments (Dikic, Wakatsuki, et Walters 2009). Regroupées dans ces 5 grandes classes, une vingtaine de familles d'UBD ont été identifiées pour jouer divers rôles dans la cellule (tableau 2).

**Tableau 2 : Les différentes classes de *Ubiquitin binding domain***

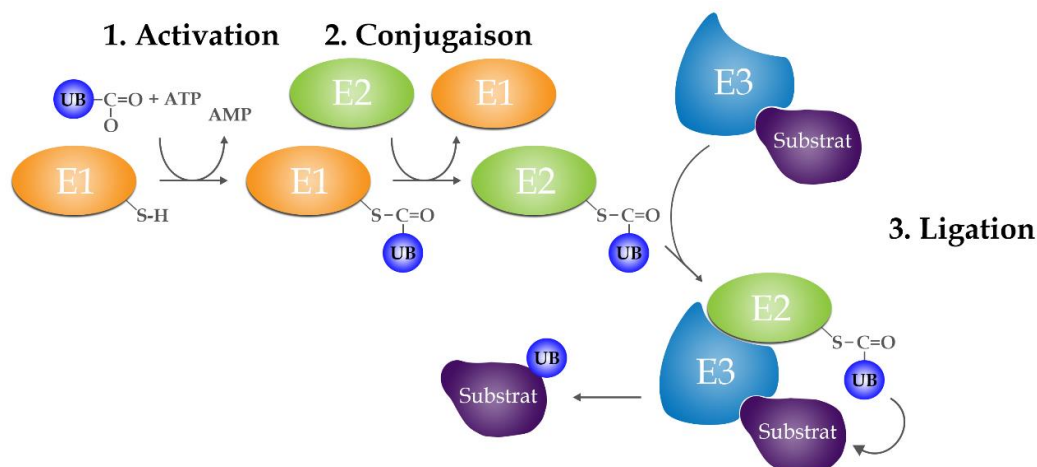
<i>Classe de UBD</i>	<i>Famille</i>	<i>Fonctions</i>
<i>Hélice-<math>\alpha</math></i>	UBA	Protéasome, autophagie et régulation des kinases
	CUE	Endocytose et régulation des kinases
	UIM	Protéasome, endocytose, réparation de l'ADN et biogenèse des corps multivésiculaires
	MIU	Endocytose
	DUIM	Biogenèse des corps multivésiculaires
	VHS	Biogenèse des corps multivésiculaires
	GAT	Biogenèse des corps multivésiculaires
	UBM	Réparation des dommages
<i>Doigt de Zinc</i>	NZF	Dégradation associée au réticulum endoplasmique, biogenèse des corps multivésiculaires et régulation des kinases
	ZnF-A20	Endocytose et régulation des kinases
	ZnF-UBP	Protéasome et autophagie
	UBZ	Réparation de l'ADN et signalisation
<i>Homologie avec Pleckstrin</i>	PRU	Protéasome
	GLUE	Biogenèse des corps multivésiculaires
<i>Relatif à la conjugaison de l'Ub</i>	UEV	Réparation des dommages, biogenèse des corps multivésiculaires et régulation des kinases
	UBC	Transfert de l'ubiquitine
<i>Autres</i>	SH3	Endocytose
	PFU	Dégradation associée au réticulum endoplasmique
	JAB1	Épissage des ARN
	UBAN	Signalisation

La liaison des UBD sur un substrat comporte trois aspects qui vont réguler la spécificité de la reconnaissance. Le premier aspect est qu'un domaine peut reconnaître l'espace entre les molécules d'ubiquitines. En effet, les différents types de chaînes d'ubiquitine adoptent des conformations ouvertes (liens poly-K63) ou encore fermées (liens poly-K48) dont la topologie est reconnue par certains UBD.

Une autre stratégie utilisée pour la reconnaissance du signal d'ubiquitine est l'orientation tridimensionnelle des ubiquitines dans l'espace. Puisque les molécules d'Ub sont reconnues via leur domaine hydrophobe, leur orientation dans les chaînes permet aux UBD de reconnaître un type de chaîne en particulier. Finalement, le substrat qui est modifié par l'ubiquitine participe aussi à la spécificité de la reconnaissance de la protéine qui contient le UBD.

## **1.7 Cascade enzymatique**

Comme mentionné précédemment, l'ubiquitine nécessite l'aide d'enzymes pour permettre son attachement sur une protéine. Cette cascade enzymatique comporte trois grandes étapes et fait intervenir le même nombre d'enzymes clés. La première étape de la réaction est celle d'activation de l'ubiquitine par l'enzyme E1. Dans un premier temps l'enzyme E1 lie l'ubiquitine ainsi que l'ATP et catalyse une réaction d'acyl-adénylation en C-terminal de l'ubiquitine (Figure 7). Par la suite, l'ubiquitine adénylée est transférée sur la cystéine catalytique de la E1 pour former un lien thioester entre les deux molécules (Schulman et Harper 2009). La seconde grande étape de la cascade enzymatique est celle de conjugaison qui fait intervenir la classe d'enzyme appelée E2. Cette étape consiste au transfert de l'ubiquitine de la E1 sur la cystéine active de la E2 dans une réaction de transtioesthérification. Finalement, la dernière étape est celle de liaison qui se résume au transfert de l'ubiquitine de la E2, sur le substrat, à l'aide d'une troisième classe d'enzyme appelée E3. Cette dernière étape peut différer entre les différentes familles de E3 puisque certaines servent seulement à la reconnaissance du substrat alors que d'autres interagissent avec l'ubiquitine avant de la transférer au substrat.



**Figure 7 : Cascade enzymatique**

L'ubiquitine s'attache de manière covalente aux protéines cibles grâce à une cascade enzymatique qui fait intervenir 3 grandes familles d'enzymes. La première étape, consiste à l'activation de l'ubiquitine via une enzyme E1 qui attaque la modification sur elle. Par la suite, l'ubiquitine sera transférée sur l'enzyme E2 lors de l'étape de conjugaison. Finalement la dernière étape est celle de ligation qui fait intervenir une enzyme E3 qui résultera en l'ubiquitinylation du substrat et la libération de celui-ci.

### 1.7.1 La famille des enzymes E1

Chez l'homme, seulement deux E1 associées à l'ubiquitine ont été identifiées, soit la protéine Uba1 qui a été pendant près d'une décennie la seule enzyme E1 connue et plus récemment, la protéine Uba6 (J. Jin et al. 2007). Bien que les deux gènes soient exprimés dans tous les tissus, Uba1 est nettement plus enrichie dans la cellule comparativement à Uba6 (X. Yang et al. 2013). En fait, Uba1 est capable d'interagir avec toutes les E2 sauf une qui est spécifique à l'enzyme Uba6, c'est entre autres pour cette raison que Uba1 est connue comme l'enzyme canonique alors que Uba6 est celle non canonique (J. Jin et al. 2007). Cette faible différence entre les deux protéines est cependant suffisante pour conférer une spécificité à chacune d'entre elles puisque ces deux enzymes ont certes des substrats en commun mais aussi des cibles variées (Xianpeng Liu et al. 2017). Une autre différence majeure entre ces deux enzymes est que Uba6 est capable d'activer FAT-10, une protéine *Ubiquitin-like* (Chiu, Sun, et Chen 2007).

L'enzyme Uba1 interagit avec l'ubiquitine via trois interfaces qui ont été caractérisées par plusieurs études de structure initialement menées chez la levure et récemment chez l'humain. La première interface regroupe les acides aminés qui font partie de la poche hydrophobe Ile44 (Leu8, Ile44, His68 et Val70) en plus de la Gly47. La deuxième interface implique les acides aminés Thr12, Asp32, Lys33 et Glu42 alors que la troisième se concentre autour des acides aminés en C-terminal de l'ubiquitine (Arg72, Leu73 et Arg74) où la réaction d'activation doit avoir lieu sur la Gly76 (Lv et al. 2018). Bien que plusieurs acides aminés ont été démontrés pour avoir une importance dans l'interaction entre l'Uba1 et l'ubiquitine, seulement la mutation de l'arginine 72 semble empêcher complètement l'activation de l'ubiquitine par la E1 alors que les autres mutants n'ont pas ou très peu d'effet sur la réaction enzymatique (Singh et al. 2017). La liaison de la E1 avec l'ubiquitine entraîne des changements de conformation dans la structure de l'enzyme exposant par le fait même son domaine d'interaction avec les E2 (I. Lee et Schindelin 2008).

### 1.7.2 La famille des enzymes E2

On dénombre chez l'humain ~40 enzymes E2 qui peuvent se diviser en 17 grandes familles selon des analyses phylogénétiques (Michelle et al. 2009). Cette classe d'enzyme se caractérise par la présence d'un domaine d'environ 150 acides aminés nommé *ubiquitin-conjugating* (UBC) qui permet l'interaction avec l'ubiquitine via la cystéine catalytique de la E2. Chacune des E2 peut interagir avec une panoplie de protéines E3 pour permettre le transfert de l'ubiquitine sur la E3-ligase (famille des HECT et RBR) ou encore directement sur le substrat (famille des RING).

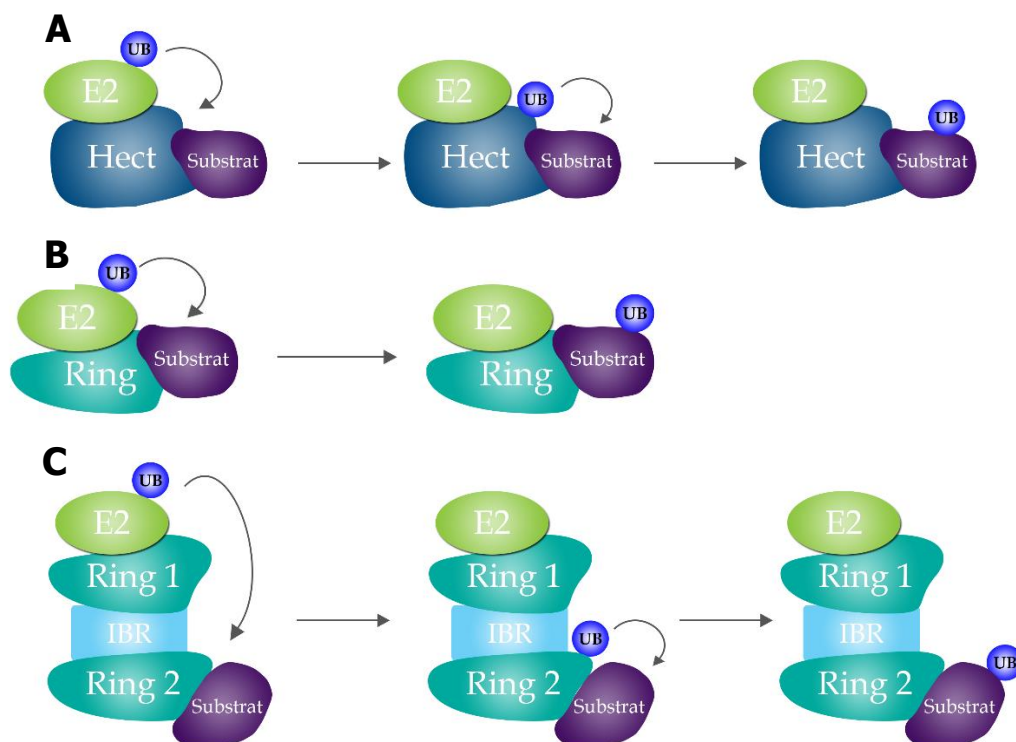
On peut séparer les E2 selon 3 catégories : (1) celles qui vont initier la chaîne (mono-ubiquitylation), (2) celles qui vont construire la chaîne (poly-ubiquitylation) et (3) celles qui peuvent faire les deux. La première catégorie regroupe les enzymes E2 qui sont capables de transférer la première molécule d'ubiquitine sur le substrat pour former une mono-ubiquitylation. Parmi celles-ci, certaines ne présentent aucune spécificité quant à la lysine à être modifiée, telles que les membres de la famille Ube2D qui vont simplement transférer l'ubiquitine sur

une lysine à proximité de leur domaine actif (Kirkpatrick et al. 2006). Cependant, d'autres E2 reconnaissent une lysine spécifique sur leur substrat tel que l'enzyme UBE2T qui transférera l'ubiquitine sur FANCD2 sur un site spécifique (Machida et al. 2006). La deuxième catégorie est celle qui forme les chaînes de poly-ubiquitines. Plusieurs études ont démontré que certaines E2 synthétisent un seul type de chaîne tel que UBE2S qui forme des chaînes K11, UBE2K qui fait des liens K48 ou encore UBE2N qui façonne uniquement des chaînes K63 (Ye et Rape 2009). D'autres E2 telles que UBE2D peuvent former différents types de liens et c'est son association avec les différentes E3 qui lui confère une spécificité dans ses liens (Kim et al. 2007). Finalement, la dernière catégorie est les E2 qui peuvent à la fois initier une chaîne et promouvoir son élongation. Seuls quelques exemples de ce type de E2 sont décrits dans la littérature et leur mécanisme d'action n'est pas encore totalement élucidé. C'est entre autres le cas pour l'enzyme E2 Cdc34 chez la levure qui peu à la fois transfère l'ubiquitine initiale sur son substrat et former des chaînes K48 par la suite. Cependant, le manque de spécificité et d'affinité entre la E2 et son substrat amène à penser que la mono-ubiquitylation est le résultat d'une collision entre les deux protéines dû au hasard plutôt qu'à un processus contrôlé (Petroski et Deshaies 2005c).

Les chaînes de poly-ubiquitine peuvent être construites grâce à l'action séquentielle de plusieurs E2 qui interagissent avec la E3. C'est le cas par exemple pour les chaînes mixtes ou branchées qui font intervenir à tour de rôle des enzymes qui synthétisent des liaisons spécifiques (Meyer et Rape 2011). Dans le cas des chaînes homogènes, la E3 peut interagir avec la même E2 pour construire une chaîne ou dans certains cas, les chaînes peuvent aussi être construites directement sur la E2 et ensuite être transférées en bloc sur le substrat (Ravid et Hochstrasser 2007). Le choix du mécanisme d'action pour la construction des chaînes dépend du rôle physiologique de celles-ci. En effet, si la protéine doit rapidement être dégradée, par exemple lors d'un mauvais repliement, l'ubiquitylation de celle-ci est faite par transfert en bloc d'une chaîne pour s'assurer d'une réaction rapide.

### 1.7.3 La famille des enzymes E3

Les E3 sont la classe d'enzymes la plus abondante dans le processus d'ubiquitylation avec plus de 1000 membres (W. Li et al. 2008a). La grande diversité de cette famille peut facilement être expliquée par le fait que ce sont ces protéines qui sont responsables de la reconnaissance des différents substrats à être ubiquitylés. Malgré ce large éventail de membres, les E3 peuvent facilement être divisées en trois grandes familles selon leur mode opératoire. La première famille à avoir été découverte est celle des *homologous to E6AP C-terminus* (HECT) (Figure 9A). Par la suite la famille des *really interesting new genes* (RING) (Figure 9B) a été caractérisée, et plus récemment, la famille des *RING-between-Ring* (RBR) (Figure 9C). Leurs principales différences se concentrent autour de leur mode opératoire pour le transfert de l'ubiquitine sur le substrat. La famille des HECT possède une cystéine catalytique qui permet le transfert de l'ubiquitine sur elle-même avant de le transférer sur le substrat. La famille des RING n'interagit jamais avec l'ubiquitine mais sert plutôt de plateforme de recrutement entre les différentes E2 et les substrats. Finalement les RBR interagissent avec la E2 via leur premier domaine RING (RING1), ce qui permet le transfert de l'ubiquitine sur leur deuxième domaine RING (RING2) et finalement au transfert de l'ubiquitine sur le substrat.



**Figure 8 : Représentation des 3 grandes familles de E3**

A) La famille des RING n'interagit jamais avec l'ubiquitine mais permet plutôt le rapprochement des substrats avec les différentes E2 pour permettre l'ubiquitinylation des protéines cibles. B) Les E3 de la famille des HECT vont dans un premier temps recruter une E2 qui transfère l'ubiquitine sur la E3. Par la suite, la E3 transfère cette dernière sur le substrat. C) La famille des RBR est la dernière famille à avoir été découverte. Les E3 de cette famille se composent d'au moins deux domaines RING sur la même protéine qui interagissent dans un premier temps avec une E2 pour permettre le transfert de l'ubiquitine sur elle-même et finalement la transférer sur le substrat via ses deux domaines.

## 1.8 Famille des RING

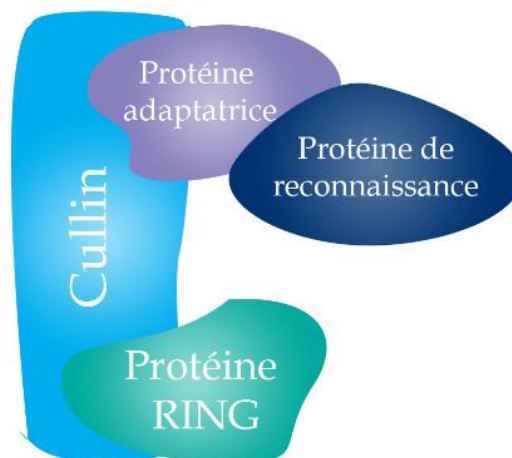
Parmi les 3 classes de E3, c'est la famille des RING qui possède le plus grand nombre de membres avec ~600 protéines (W. Li et al. 2008b). Le domaine RING se compose d'un motif doigt de zinc spécifique à cette famille qui implique la liaison avec deux ions de zinc. D'autres protéines de la famille des RING possèdent un domaine U-box qui est en tout point ressemblant au domaine classique mais qui ne fait pas intervenir d'ions zincs (Ohi et al. 2003). Leur domaine de reconnaissance au substrat ainsi que leur activité catalytique peut se retrouver sur une même protéine



(exemple de c-CBL) ou encore se diviser sur plusieurs protéines telles que pour le complexe APC ou encore les complexes *Cullin-RING E3 ligases* CRLs.

### 1.8.1 Complexe CRLs

Parmi les membres de la famille des RING, la classe des CRLs est la plus nombreuse avec la formation de plus de 200 complexes qui seraient responsables d'environ 20% de l'ubiquitinylation total dans la cellule (Petroski et Deshaies 2005a). Tous les complexes CRLs se composent d'une protéine qui contient le domaine RING, d'une Cullin, d'une protéine adaptatrice, et finalement, une protéine responsable de la reconnaissance des substrats (Figure 9). La Cullin sert de protéine d'échafaudage au complexe en permettant de lier et de rapprocher la protéine responsable de la reconnaissance de la E2 (celle qui comporte le domaine RING) et le module de reconnaissance des substrats. Chez l'homme, on dénombre 8 complexes CRLs (CRL1, CRL2, CRL3, CRL4A, CRL4B, CRL5, CRL7 et CRL9) qui sont nommés en fonction de la Cullin qu'ils contiennent.



**Figure 9 : Composition des complexes CRLs**

Tous les complexes CRLs partagent la même composition protéique avec la présence d'une protéine comportant le domaine RING (ROC1 ou ROC2) qui permet l'interaction avec la protéine E2. Un des membres de la famille des Cullins, une protéine adaptatrice (exception faite pour le complexe CRL3) et une protéine responsable de la reconnaissance des substrats.

### 1.8.1.1 Protéine RBX

Dans les complexes CRLs, ce sont les protéines ROC1/RBX1/Hrt1 ou ROC2/RBX2/SAG qui possèdent le domaine RING et qui sont donc capables d'interagir avec les différentes enzymes E2 pour permettre le transfert de l'ubiquitine sur le substrat (Ohta et al. 1999). Les deux membres de la famille sont exprimés chez l'humain et sont conservés chez tous les mammifères (Sun et al. 2001). Bien que certains complexes présentent une préférence entre les deux protéines, la majeure différence entre elles est que ROC1 est exprimée de manière constitutive dans les cellules alors que la protéine ROC2 est exprimée suite à divers stress par la présence de plusieurs sites consensus pour différents facteurs de transcription tels que NFκB, c-MYC, p53 et plusieurs autres (Wingender et al. 1996). Bien que *in vitro*, les deux membres de la famille RBX peuvent interagir avec toutes les cullins, il semblerait que les complexes interagissent préférentiellement avec l'une ou l'autre des protéines *in vivo* (Ohta et al. 1999) (X. Wang et al. 2015). Un autre rôle important pour les protéines RBX est que celles-ci servent d'enzyme E3 pour la modification *ubiquitin-like* NEDD8. Lorsque les protéines RBX1 interagissent avec les cullins, ils catalysent la NEDDylation de celles-ci, ce qui a comme conséquence l'activation des complexes CRLs (Huang et al. 2009).

### 1.8.1.2 Cullin

Les membres de la famille des cullins présentent tous une structure similaire avec deux régions caractéristiques, la première en N-terminal se compose d'une région riche en hélice-α appelées NTD (*N-terminal domain*) et une deuxième qui adopte une structure globulaire nommée CTD (*C-terminal domain*). Bien que les protéines de la famille des cullins ne possèdent pas d'activité enzymatique, celles-ci se retrouvent au centre du complexe CRL et permettent l'organisation de celui-ci par ses interactions avec ces diverses composantes. Le domaine NTD est responsable de la liaison avec la protéine adaptatrice et se compose de trois répétitions du domaine Cullin qui est composé de deux hélices courtes suivies de trois longues hélices (Zheng et al. 2002). Le domaine CTD est composé de quatre

structures différentes qui adoptent une conformation globulaire en plus de contenir un domaine d'homologie avec les Cullin. (Zheng et al. 2002). Cette région de la protéine est responsable de l'interaction avec les protéines de la famille RBX qui contiennent le motif RBX.

### 1.8.1.3 Protéines adaptatrices et protéines de reconnaissance

Les protéines adaptatrices sont différentes pour chaque complexe CRL tel qu'illustré par la Figure 10, hormis pour CRL2 et CRL5 qui partagent les adaptateurs Elongin C/B. Les protéines adaptatrices interagissent avec la Cullin sur sa région NTD via des domaines SKP1/BTB/POZ pour CRL1, CRL2, CRL5 et CRL7, BPB pour CRL4 et BTB pour CRL3 (Sarikas, Hartmann, et Pan 2011). Le rôle principal de ces protéines est de recruter les différentes protéines de reconnaissance des substrats qui sont spécifiques à chacun des complexes. On dénombre près de 400 protéines de reconnaissance pour l'ensemble des complexes avec dans chaque famille entre 60 et 180 membres (exception faite pour CRL7 qui interagit seulement avec FBXW8) (Dias et al. 2002).

ROC1	Cullin 1	SKP1	F-box
ROC1	Cullin 2	Elongin C Elongin B	VHL-box
ROC1	Cullin 3		BTB protein
ROC1	Cullin 4a	DDB1	DCAF
ROC1	Cullin 4b	DDB1	DCAF
ROC2	Cullin 5	Elongin C Elongin B	SOCS-box
ROC1	Cullin 7	SKP1	FBXW8
?	Cullin 9	?	?

**Figure 10 : Composition des différents complexes Cullin**

Tous les complexes CRLs partagent la même composition protéique avec la présence d'une protéine comportant le domaine RING (ROC1 ou ROC2) qui va permettre l'interaction avec la protéine E2, un des membres de la famille des Cullins qui va interagir avec une protéine adaptatrice (exception faite pour le complexe CRL3) et une protéine responsable de la reconnaissance des substrats.

## 1.8.2 Régulation des complexes CRL

Considérant le rôle important de ces complexes protéines dans de nombreux processus cellulaires, leur activation aussi bien que leur inhibition doit être un processus strictement régulé. Les complexes CRLs possèdent des régulations qui leur sont propre à chacune. Néanmoins, certains mécanismes sont communs à tous les complexes formés par des cullins.

### 1.8.2.1 NEDDylation

Tel que mentionné précédemment, la modification *ubiquitin-like* NEDD8 a été initialement identifiée pour son rôle dans la régulation des cullins (Liakopoulos et al. 1998). Cette modification prend place dans le domaine CTD de ces dernières et amène un changement de conformation qui permet aux protéines RBX d'approcher la E2 plus près du substrat pour permettre le transfert de l'ubiquitine sur celui-ci (Duda et al. 2008). La de-NEDDylation des cullins se fait par un complexe appelé CSN (COP9 Signalosome) qui suite à son interaction avec un complexe CRL-NEDD8, change de conformation afin d'exposer son site catalytique pour cliver NEDD8 (Lingaraju et al. 2014).

### 1.8.2.2 CAND1

CAND1, pour Cullin-associated NEDD8-dissociated factor 1, a initialement été rapporté pour son rôle dans l'inhibition des complexes CRLs (J. Liu et al. 2002). Des études de structures ont démontré que CAND1 liait la partie CTD et NTD des Cullin non-NEDDylée pour empêcher leur modification mais aussi la liaison des protéines de reconnaissance des substrats (Goldenberg et al. 2004). Cependant des études génétiques ont démontré qu'en fait CAND1 pourrait promouvoir l'activité des complexes de Cullin (W. Zhang et al. 2008). Les deux hypothèses contradictoires ont été réconciliées en 2013 dans une étude de structure qui a démontré que CAND1 permettait en fait la dissociation des protéines de reconnaissance des substrats et permettrait l'échange entre ces différentes protéines sur un complexe CRL (Pierce et al. 2013).

La régulation des complexes CRLs oscille entre des états actifs et inactifs dépendamment de leur interaction avec les différents régulateurs positifs et négatifs. Un modèle de régulation a été proposé où dans un premier temps le complexe se forme avec une protéine de reconnaissance des substrats. Par la suite, ce complexe peut être soit NEDDylé pour activer le processus d'ubiquitinylation, soit interagir avec CAND1 pour permettre le changement de protéine de reconnaissance. Lorsque la quantité de substrats va diminuer, le complexe CSN va dé-NEDDylé le complexe, pour que celui puisse retourner à l'état initial et rentrer dans le cycle à nouveau.

## **1.9 Complexe CRL4**

Le complexe CRL4 a comme particularité qu'il est le seul membre de cette famille utilisant deux différentes Cullin (Cullin 4A et 4B) pour former un complexe fonctionnel. Mis à part les deux protéines Cullin4, celui-ci se compose de la protéine adaptatrice DDB1 (*DNA damage-binding protein 1*) et des protéines de reconnaissance des substrats, les DCAFs (DDB1-Cullin4 associated factor) (Figure 10).

### **1.9.1 Cul4A et Cul4B**

Les protéines Cul4A et Cul4B sont des paralogues qui partagent plus de 80% d'homologie de séquence entre elles avec comme principale différence une extension de 149 acides aminés en N-terminal pour Cul4B (Zou et al. 2009). Cette séquence en plus contient un NLS (*Nuclear Localization Signal*) ce qui permet à la protéine Cul4B d'être nucléaire alors que la protéine Cul4A est majoritairement cytoplasmique avec une sous-population qui est nucléaire. Les deux protéines sont présentes dans de nombreux tissus avec différents niveaux d'expression d'après plusieurs études à large échelle (Su et al. 2004; Uhlén et al. 2015). Cependant, aucune étude ne s'est intéressée de manière plus spécifique à leur expression respective.

En ce qui regarde leur fonction distinctive, plusieurs études ont démontré que les deux protéines possédaient des fonctions redondantes. En effet, tout dépendant de la cible évaluée ou du mécanisme cellulaire à l'étude, les deux protéines peuvent jouer un rôle similaire. C'est le cas entre autres pour la protéine Cdt1 qui est dégradée par le complexe DDB1-Cul4. L'utilisation d'ARN interférant a démontré que les 2 protéines doivent être diminuées pour que la protéine Cdt1 se stabilise (Senga et al. 2006). Néanmoins, les deux protéines présentent des fonctions propres à chacune puisque lors de la délétion d'une des deux protéines, les cellules présentent des phénotypes apparents et qui ne sont donc pas compensés par la présence du second homologue (Hannah et Zhou 2015).

### **1.9.2 DDB1**

La protéine DDB1 a initialement été découverte en association avec la protéine DDB2 pour être défectueuse dans la maladie *Xeroderma pigmentosum* (Patterson et Chu 1989). Cette maladie se caractérise entre autres par un défaut dans les mécanismes de réparation de l'ADN suite aux dommages aux UVs, ce qui se traduit par une augmentation drastique de l'incidence du cancer chez ces patients. Dans l'une des formes de cette maladie, la protéine DDB2 est mutée et est incapable de lier l'ADN, ce qui empêche donc la réparation (Tang et Chu 2002). Ainsi, les protéines DDB1-DDB2 ont été initialement caractérisées pour être impliquées dans la réparation des dommages à l'ADN (R. T. Johnson et al. 1985). C'est quelques années plus tard que la fonction d'ubiquitylation du complexe DDB1-DDB2 a été élucidée entre autres par la découverte de son interaction avec la protéine Cullin4 (Shiyanov, Nag, et Raychaudhuri 1999). Par la suite, de plus en plus d'études ont commencé à rapporter la présence de DDB1 avec la protéine Cullin4 en absence de DDB2. De plus, l'identification de la composition des autres complexes CRLs et la mise en évidence des protéines adaptatrices dans ceux-ci ont été la prémisse à l'identification du rôle principal pour DDB1 au sein du complexe CRL4 (Hu et al. 2004).

Des homologues de la protéine DDB1 ont été identifiés chez de nombreux eucaryotes et celle-ci est très bien conservée chez les mammifères (Fu et al. 2003; Wittschieben et Wood 2003). Au niveau de sa structure, DDB1 possède 3  $\beta$ -propeller (BP) nommé BPA, BPB et BPC en plus d'un domaine en C-terminal (CTD) (Scrima et al. 2008). Chacune des structures BP se compose de 7 domaines WD40 qui sont des domaines utilisés pour la reconnaissance entre protéines ou encore entre protéine et ADN. C'est d'ailleurs par ces domaines que DDB1 reconnaît les différentes protéines de reconnaissance de substrats et les deux Cullin4. Mis à part son rôle comme protéine adaptatrice dans le complexe CRL4, DDB1 ne possède aucune autre fonction connue à ce jour et bien qu'elle ne possède aucune activité enzymatique, son rôle de protéine adaptatrice est suffisant pour la rendre essentielle à la survie cellulaire puisque la délétion de celle-ci est létale à l'état embryonnaire (Cang et al. 2006).

### **1.9.3 DCAFs**

La famille des DCAFs compte près de 60 protéines qui ont été, pour la grande majorité, identifiées par deux groupes de chercheurs la même année grâce à des études de spectrométrie de masse et de bio-informatique (Angers et al. 2006; He et al. 2006). En effectuant des immunoprécipitations (IPs) de la protéine DDB1, ils ont été en mesure d'identifier des protéines qui se retrouvaient en complexe avec celle-ci ainsi qu'avec les protéines Cullin4A et B. Par cette approche, ils ont identifié près d'une vingtaine de DCAFs dont certaines étaient déjà bien connues. De plus, en comparant la séquence de celles-ci, ils ont été en mesure d'identifier un domaine WD40 consensus responsable de l'interaction avec la protéine DDB1. Ainsi, en se basant sur la présence de ce domaine, plusieurs protéines contenant des domaines similaires ont été associées à cette famille de protéines (J. Lee et Zhou 2007a). Bien que certaines de ces protéines ont déjà démontré leur fonction cellulaire comme protéine adaptatrice du complexe CRL4, beaucoup d'entre elles restent néanmoins peu caractérisées et sont restées associées à cette famille de protéines par la seule présence d'un domaine consensus. De plus, certaines des DCAF, tel que la protéine

DDA1, ne présentent pas de domaine WD40 et est toutefois capables de former un complexe fonctionnel avec DDB1-Cul4 pour l'ubiquitinylation de ses substrats (Shabek et al. 2018). Ceci suggère que la liste des protéines appartenant à la famille des DCAF n'est pas exhaustive et que des études plus rigoureuses doivent être effectuées.

## **1.10 Hypothèse et objectifs**

### **1.10.1 Chapitre 1 : Étude des cibles du complexe DDB1-Cul4 par spectrométrie de masse via l'intermédiaire des protéines DCAFs**

Le premier chapitre découle de précédents travaux que nous avons effectués dans le laboratoire et qui ont orienté nos recherches sur le complexe d'ubiquitine ligase DDB1-Cul4 (Dubois et al. 2016). Dans ce précédent article, nous avons découvert que le complexe DDB1-Cul4 se retrouvait dans l'environnement du complexe MCM qui est impliqué au niveau de la réplication de l'ADN et cette association semble augmenter en réponse aux dommages à l'ADN. Bien que l'implication du complexe DDB1-Cul4 dans la réplication et la réparation de l'ADN aient déjà été rapportées, la E3 responsable de l'ubiquitinylation du complexe MCM n'a pas encore été découverte. Tel que mentionné, le complexe DDB1-Cul4 utilise les protéines DCAFs pour reconnaître ses différents substrats. Cette famille de protéines regroupe près de 60 membres qui ont été pour la majorité identifiées par des approches de spectrométrie de masse et de bio-informatique. Ainsi, si certaines sont déjà bien caractérisées, il reste cependant une portion des protéines qui n'ont pas été sujettes à de plus amples investigations. En conséquence, il y a des protéines DCAFS pour lesquelles très peu voir aucune cible où fonction n'ont été rapportées ou encore, des cibles du complexe qui ne sont pas associées a de protéines DCAFs (tel que les MCM). C'est pour ces raisons que nous avons décidé d'investiguer les fonctions du complexe CRL4.



**Hypothèse : L'utilisation des protéines DCAF nous permettra d'identifier les cibles et les fonctions du complexe DDB1-Cul4.**

Objectif 1 : Identifier les protéines qui agissent comme protéines de reconnaissance des substrats pour le complexe DDB1-Cul4.

Objectif 2 : À l'aide de différentes approches de spectrométrie de masse, caractériser chacune des protéines DCAFs.

**1.10.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène**

Le second chapitre présente des travaux effectués sur le pseudogène de l'ubiquitine UBBP4. Initialement rapportée comme un pseudogène, l'équipe du Pr Roucou a découvert des peptides uniques issus du gène de UBBP4 lors de la réanalyse de données de spectrométrie de masse. Cette découverte, qui avait initialement comme objectif de trouver une fonction aux protéines alternatives via l'identification de leur partenaire d'interaction, a permis d'associer UBBP4 à des fonctions liées à l'ubiquitine. La protéine alternative identifiée vient en fait d'un second cadre de lecture que celui qui a été annoté initialement sur le gène UBBP4. Nous avons donc commencé à investiguer les informations précédemment rapportées pour ce pseudogène, pour nous rendre compte que celui-ci avait initialement été mal annoté dû à une erreur de séquençage (détaillés dans les résultats). Nous avons conclu des résultats préliminaires que UBBP4 n'était pas un pseudogène de l'ubiquitine mais en fait un paralogue de celui-ci et qu'il pouvait produire quatre variants de l'ubiquitine dans deux cadres de lectures différents.

**Hypothèse : Le pseudogène UBBP4 n'est pas un pseudogène et celui-ci peut produire des variants de l'ubiquitine**

Objectif 1 : Définir les variants d'ubiquitine qui sont exprimés à partir du gène UBBP4.

Objectif 2 : Préciser le rôle de Ub<sup>KEKS</sup> dans les fonctions cellulaires.

## **2. Matériel et méthode**

### **2.1 Général**

#### **2.1.1 Analyse de l'expression des ARNm à partir d'une base de données**

Les données d'expression des ARNm ont été obtenues à partir de la base de données Expression Atlas (Petryszak et al. 2016). Pour le chapitre 1, les données d'expression de l'ARNm proviennent exclusivement du GTEx projet qui regroupe l'expression de plusieurs tissus humains (Lonsdale et al. 2013). Le niveau des ARNm est exprimé en TPM (transcripts per million). Pour le chapitre 2, 5 études ont été utilisées pour identifier l'expression en ARNm de l'ubiquitine ainsi que de ses 5 pseudogènes. Puisque ces 6 gènes produisent des ARNm qui peuvent être communs, nous nous sommes assuré que pour chacune de ces études les fragments séquencés pouvant aller sur plusieurs gènes ont été pairés seulement avec une seule région du génome afin d'éviter les duplications de signal. Les résultats ont été exprimés en TPM X100.

#### **2.1.2 Culture cellulaire, transfection et milieux SILAC**

Toutes les cellules utilisées dans ce travail ont été cultivées dans du milieu DMEM (*Dulbecco's modified eagle medium*) (Invitrogen Life Technologies) supplémenté avec 10% de FBS (*Fetal Bovin Serum*) (Wisent). Les transfections transitoires et stables ont été effectuées avec la Lipofectamine LTX en suivant les indications du fournisseur. Pour les lignées stables, les cellules ont été sélectionnées avec 5µg/ml de blasticidine (Wisent) et 10µg/ml d'hygromycine (Wisent). Les inductions ont été faites avec 2µg/ml de doxycycline pour le temps indiqué pour chacune des expériences (Clontech).

Les milieux SILAC ont été obtenus en utilisant du milieu DMEM sans lysine et arginine (Invitrogen Life Technologies) contenant 10% de FBS triplement dialysé (Invitrogen Life Technologies) supplémenté avec de la lysine ou de l'arginine isotopiques soit

légers (Arg0, Lys0), moyens (Arg6, Lys4) ou lourds (Arg10, Lys8) (Cambridge Isotope Laboratories). La concentration finale de l'arginine était de 28pg/ml et de 49pg/ml pour la lysine. Les cellules ont été cultivées dans les milieux respectifs pour un minimum de 6 passages afin de s'assurer d'une incorporation optimale des isotopes des deux acides aminés.

### **2.1.3 Extraits cellulaires et immunobuvardage**

Pour faire les extraits totaux, les cellules ont été décollées avec de la trypsine 0,25% à 37°C pendant 10 minutes, centrifugées, rincées avec du PBS1X (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, et un pH de 7.4) et recentrifugées. Par la suite, elles ont été directement lysées dans du Laemli 1X préparé à partir du 5X (50% glycérol, 10% SDS, 300 mM Tris pH 6.8) sans bleu de bromophénol et les extraits ont été dosés en utilisant la trousse Pierce BCA (Thermo Scientific). Les extraits ont été séparés sur gel SDS de polyacrylamide. Les protéines ont été transférées par la suite sur une membrane de nitrocellulose (PerkinElmer) et pour les profils d'ubiquitine, la membrane était par la suite bouillie pendant 10 minutes dans de l'eau. Le blocage était effectué dans du TBS-T (10mM Tris pH 7.4, 150 mM NaCl, 0.1% Tween-20) 5% lait pendant 1h. Les anticorps ont été dilués dans la même solution que le blocage à différentes concentrations (tableau 3) et incubé toute la nuit à 4°C. Les membranes ont été lavées un minimum 3 fois dans du TBS-T à température ambiante et incubée 2h avec l'anticorps secondaire approprié (tableau 3) dans le même tampon que le blocage. Après le temps d'incubation, les membranes ont été lavées 3 fois avec le TBS-T et révélées à l'aide du Clarity Western ECL (Bio-Rad Laboratories).

**Tableau 3 : Anticorps primaires et secondaires utilisés**

<b>Anticorps primaires</b>	<b>Dilution</b>	<b>Compagnie</b>	<b>Référence</b>
GFP	1/3000	Roche	#11814460001
HA	1/10000	Invitrogen	#26183
Actine	1/25000	Sigma	#A5441
TCERG1	1/1000	Bethyl	#A300-360A
LMNA	1/2000	Abcam	#ab133256
LMNB2	1/1000	Cell Signaling	#12255
Nucleolin	1/5000	Abcam	#ab136649
<b>Anticorps secondaires</b>	<b>Dilution</b>	<b>Compagnie</b>	<b>Référence</b>
Souris-HRP	1/10000	Cell Signaling	#7076S
Lapin-HRP	1/10000	Cell Signaling	#7074S
Lapin Alexa Fluor 546	1/500	Invitrogen	A11010
Souris Alexa Fluor 488	1/500	Invitrogen	A11001

### 2.1.4 Immunofluorescence

Pour les immunofluorescences, 15 000 cellules ont étéensemencées sur des lamelles dans un 24 puits et laissées pour 24h. Les cellules ont été lavées 3 fois au PBS 1X froid et fixées au paraformaldéhyde 4% dans du PBS pour 20 minutes. Après 2 lavages au PBS 1X, les cellules ont été perméabilisées en utilisant 0.15% de triton X-100 dans du PBS 1X pour 5 minutes et après 2 lavages, les lamelles ont été bloquées en utilisant 10% de sérum de chèvre dans du PBS 1X pour 20 minutes. Les anticorps primaires ont été dilués dans du PBS 1X supplémenté ave 10% sérum de chèvre et les lamelles ont été incubées 16h pour LMNA et pour HA ou 2h pour la nucleolin (tableau 3) à température de la pièce. Les cellules ont par la suite été lavées 2 fois avec du PSB 1X et incubées avec l'anticorps secondaire (tableau 3) approprié pour 1h dans du PBS1X avec 10% sérum de chèvre. À la suite des deux lavages, les noyaux ont été marqués au DAPI (1µg/µl) pour 10 minutes dans du PBS 1X et lavés deux fois par la suite. Finalement, les lamelles ont été montées en utilisant du liquide de montage immuno-mount (ThermoFisher Scientific).

### **2.1.5 Digestion des protéines**

#### En gel

Les protéines contenues dans les IPs ont été réduites avec 10mM de DTT et bouillies à 95°C pendant 5 minutes avant d'être alkylées dans 50mM d'iodoacétamide et séparées sur un gel d'acrylamide gradient 4-12% (NuPAGE, Thermo Fisher Scientific). Les gels ont été colorés au SimplyBlue SafeStain (Invitrogen Life Technologies) pour être visualisés avant que les IPs soient découpées en huit bandes et sujettes à une digestion à la trypsine (Trypsin Gold, Promega Corporation) (Shevchenko et al. 1996). Les peptides résultants de la digestion ont été extraits avec 100% acétonitrile et 1% d'acide formique. Les peptides ont été déshydratés en utilisant un speedvac et réhydratés dans 0.1% TFA et le protocole de Zip tip® (Zip tip® Pipette Tips, EMD Millipore Corporation) a été utilisé pour enlever toutes traces de détergents et de sels dans les échantillons.

#### Sur billes

Les IPs ont été lavées 5 fois avec une solution de bicarbonate d'ammonium 20mM et incubées 30 minutes à 60°C en présence de 10mM de DTT dans le tampon d'ammonium. Par la suite, 15mM d'iodoacétamide a été ajouté et les IPs ont été incubées 1h dans le noir. La solution d'iodoacétamide a été neutralisée par l'ajout de DTT 1M pour atteindre une concentration finale de 15mM de DTT et laissée 15 minutes à température pièce. La trypsine a été ajoutée directement sur les billes à 37°C pour une durée de 16h. La digestion a été arrêtée par l'ajout de 1% acide formique et le surnageant a été transféré dans un nouvel ependorf. Les billes ont été lavées avec une solution de 60% d'acétonitrile/0.1% d'acide formique et les lavages ont été ajoutés au premier surnageant. Les échantillons ont été asséchés et réhydratés dans 25µl de 0.1% TFA et sujets au protocole de Zip tip® (Zip tip® Pipette Tips, EMD Millipore Corporation).

### **2.1.6 Spectrométrie de masse**

Les peptides ont été séparés par le système Dionex Ultimate 3000 nanoHPLC. 12  $\mu$ l d'échantillon (pour un total de 2  $\mu$ g de peptides) a été injecté avec un débit de 4  $\mu$ l/min sur une colonne Acclaim PepMap100 C18 (0.3 mm id x 5 mm, Dionex Corporation, Thermo Scientific). Après leur concentration, les peptides ont par la suite été élués sur une colonne PepMap (75  $\mu$ m x 50 cm, Dionex Corporation Thermo Scientific) avec un gradient linéaire de 5-35% de solvant B (90% d'acétonitrile et 0.1% d'acide formique) pour une durée totale de 4h avec un débit constant de 200nl/min. Ce système de séparation est couplé à un spectromètre de masse Orbitrap QExactive (Thermo Fischer Scientific Inc.) qui utilise comme source d'ionisation un électrospray avec un voltage de 2.0 kV. Une acquisition de spectre complet (m/z 350-1600) a été effectuée dans l'orbitrap avec une résolution de 70 000 après l'accumulation de 1 000 000 d'ions. Les peptides sélectionnés pour la fragmentation par collision étaient ceux qui présentaient la plus forte intensité lors de la première acquisition. L'énergie de collision était normalisée à 35% et la résolution pour le MS/MS était réglée à 17 500 pour 50 000 ions avec un temps maximal de 250ms pour les balayages MS et de 60ms pour les MS/MS. Tous les états de charges non assignés ainsi que les ions avec des charges de 7 et 8 ont été rejetés. La liste d'exclusion dynamique a été limitée à un maximum de 500 entrées avec un temps maximal de rétention de 40 secondes et une fenêtre de 10ppm de masse relative. Pour permettre une meilleure précision lors de l'acquisition de la masse, l'option de verrouillage de la masse a été activée. Finalement, les données ont été obtenues grâce à l'utilisation du logiciel Xcalibur version 2.2 SP1.48.

### **2.1.7 Quantification et analyse bio-informatique**

L'identification et la quantification des protéines identifiées par spectrométrie de masse ont été effectuées par le logiciel MaxQuant version 1.5.2.8. Les analyses ont été faites pour du milieu SILAC avec 3 conditions (Light labels = 0, Medium labels

= Lys4 et Arg6 et Heavy labels = Lys8 et Arg10), un maximum de 2 mauvaises coupures par peptide, la trypsine comme mode de digestion, la carbamidométhylation sur cystéine a été cochée comme modification fixe et l'oxydation des méthionines et acétylation des protéines en N-terminal ont été sélectionnées comme modifications variables. La requantification a été sélectionnée pour calculer le ratio isotopique lorsque les paires d'isotopes ne sont pas identifiées (problème fréquent lors des immunoprécipitations par la faute de l'absence de la protéine dans les contrôles). La longueur minimale des peptides considérée pour la quantification était de 7 acides aminés et le FDR (false discovery rate) a été réglé pour 5%. Un minimum d'un peptide a été utilisé pour l'identification mais seulement les protéines avec deux peptides ont été sélectionnées pour les analyses. La quantification des protéines a été faite en utilisant les peptides uniques et les razors.

## **2.2 Spécifique au Chapitre 1**

### **2.2.1 Clonage et plasmide des protéines DCAFs**

Toutes les séquences des protéines DCAFs ont été obtenues de cellules U2OS à partir d'une librairie d'ADNc qui a été générée à la suite d'une RT-PCR avec un oligo-dT sur les ARNm provenant d'une extraction au TRIzol® (Invitrogen Life Technologies). La séquence de chacune des protéines DCAFs a été amplifiée en utilisant des oligonucléotides spécifiques pour chacune d'entre elles en plus de contenir les séquences spécifiques ATTB. Ces séquences permettent la recombinaison des produits de PCR dans le plasmide pDONR221 (Invitrogen Life Technologies, Burlington, ON, Canada) à l'aide d'une réaction impliquant la BP clonase II (Invitrogen Life Technologies). À la suite de la confirmation des plasmides par séquençage, les DCAFs ont été insérées dans le plasmide pbLAP1, qui permet la production de protéine de fusion avec BirA\* et l'insertion dans le génome à l'aide du système Flip-In.

### 2.2.2 Co-immunoprécipitations pour DDB1-GFP

Les cellules ont été décollées en grattant les pétris dans du PBS1X froid et lysées dans du tampon Hight-Salt (HS) (1% NP-40, 50 mM Tris pH7.5, 300 mM NaCl, 150 mM KCl, 5 mM EDTA, 1 mM DTT, 10% glycerol, *Roche Complete Protease Inhibitor Cocktail*) sur glace pendant 10 minutes. Les extraits étaient par la suite soniqués 3 fois 10 secondes à 20% d'intensité avec un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific) et laissés un autre 10 minutes sur glace. Par la suite, les lysats étaient clarifiés par centrifugation à 13 000g à 4°C et incubés avec des billes GFP (GFP-Trap-A, ChromaTek) pour 2h à 4°C. Après le temps d'incubation, les billes étaient lavées 3 fois avec du tampon HS et diluées dans du Laemli 1X.

### 2.2.3 Purification par affinité à l'aide de BirA\*

Pour chacune des expériences, deux pétris de 150mm de cellules U2OS Flip-In ont été utilisés et induits pendant 24h à la doxycycline. La biotine a été ajoutée au milieu de culture à une concentration finale de 50uM pour une période de 24h. Par la suite, les cellules ont été lysées en utilisant du tampon RIPA (50 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1,5 mM MgCl<sub>2</sub>, 0,1 % SDS, 1 % IGEPAL CA-630, 1 mM PMSF, 0,4 % désoxycholate de sodium et 1 mM DTT) et les extraits cellulaires ont été placés sur rotateur à 4°C pour une durée de 20 minutes. Puis les échantillons ont été soniqués avec une amplitude de 30% pour 3 répétitions pendant 10 secondes et replacés en rotation pour une durée de 1h à 4 °C. Par la suite, du SDS a été ajouté aux lysats cellulaires à raison d'une concentration finale de 0,4% et remis en rotation pour une durée de 15 minutes supplémentaires à la même température. Finalement, les échantillons ont été clarifiés par centrifugation à 20 000 g pour une durée de 20 minutes.

Pour la purification par affinité, les extraits ont été mis en contact avec des billes de Sepharose couplées à la streptavidine (GE Healthcare Life Sciences, Marlborough, États-Unis) préalablement lavées avec du tampon RIPA. Les billes ont par la suite



été mises en contact avec les extraits cellulaires et agitées par rotation à 4°C pour une durée de 3h. Les billes ont été lavées une fois avec du tampon de lavage BioID (2% SDS, 50 mM Tris-HCl, pH 7,5) et 3 fois avec du tampon RIPA. Finalement, les billes ont été sujettes à une digestion par la trypsine (protocole de digestion sur billes) et les peptides résultants ont été analysés par spectrométrie de masse.

#### **2.2.4 Pulse SILAC et digestion en solution**

Deux pétris de 35mm contenant les cellules exprimant les DCAFs ont été cultivés dans du milieu R0K0 jusqu'à obtention d'une confluence d'environ 60%. Au temps 0, l'un des pétris a été induit à la doxycycline pour une durée de 8h et à la suite de cette période de temps, le milieu de culture des cellules a été changé pour du milieu moyen (R6K4) pour les cellules non induites et pour du milieu lourd (R10K8) pour les cellules induites. Au temps 24h, les cellules ont été lysées directement dans 8M urée 10mM Hepes pH8,0 et soniqué 3 fois 15 secondes. Après centrifugation, les surnageants sont dosés à l'aide de la trousse Pierce BCA (Thermo Scientific) et la même quantité d'extraits cellulaires ont été mélangé ensemble. Par la suite les extraits ont été réduits à l'aide de 1µg de DTT et bouillis à 95°C pendant 2 minutes et laissés 30 minutes à température pièce. Par la suite, 5µg d'iodoacétamide ont été ajoutés et incubée à température pièce pendant 20 minutes dans le noir et finalement l'échantillon a été dilué en utilisant un tampon de 50mM d'ammonium bicarbonate pour amener la concentration de l'urée à 2M. Finalement, 50ng de trypsine est ajouté et laissé à 37°C pour une durée de 16h et à la suite de ce délai, l'échantillon a été acidifié avec du TFA pour l'amener à une concentration finale de 0.1% et a été sujet au protocole de Zip tip® (Zip tip® Pipette Tips, EMD Millipore Corporation).

## 2.3 Spécifique au Chapitre 2

### 2.3.1 Séquençage de UBBP4 pour l'ADN génomique et de l'ARNm

L'ADN génomique a été extrait à l'aide de *GenElute Mammalian Genomic DNA Miniprep Kit* (Sigma-Aldrich) à partir d'un pètri de 100mm pour chacune des lignées cellulaires (HeLA, 293 et U2OS). La région du génome correspondant spécifiquement à UBBP4 a été amplifiée par PCR en utilisant une paire d'oligonucléotides dont le premier s'apparie dans le seul intron du gène et le second dans le 2<sup>e</sup> exon. L'ARNm quant à lui a été extrait à l'aide du réactif TRIzol® (Invitrogen Life Technologies) à partir d'un pètri de 100 mm des mêmes lignées que pour l'ADN génomique. Les ADNc ont été obtenus par la suite en suivant le protocole de la ProtoScript® II *Reverse Transcriptase* (New England Biolabs) à partir de 1 µg d'ARNm en utilisant un oligo-dT. Finalement, une réaction de PCR a été faite afin d'amplifier la séquence correspondant à l'ARNm de UBBP4 comprenant les deux exons.

Pour donner suite à leur amplification, les produits de PCR (aussi bien pour l'ARNm que l'ADN génomique) ont été séparés sur un gel d'agarose 1% et les produits ont été extraits du gel avec la trousse de *Gel/PCR DNA Fragments Extraction Kit* (FroggaBio). Les produits ont par la suite été clonés dans le vecteur pUC19 à l'aide d'une digestion franche par SmaI. Grâce à une sélection bleu-blanc pour les colonies à la suite de la digestion du produit x-galactosidase, 10 clones ont été sélectionnés et envoyés pour le séquençage (Plateforme de séquençage de l'Université Laval).

**Tableau 4 : Amorces utilisées pour le séquençage de l'ARNm et du gène de UBBP4**

<b>UBBP4</b>	Génomique	ATCTAGGTCAAAATGCGGATCTTCGT
<b>UBBP4</b>	Génomique	ATGACAAAACACCAAGTATGCTACCATT
<b>UBBP4</b>	ARNm	GCTGGTGCTGCAAGAAAGTT
<b>UBBP4</b>	ARNm	TGCTACCATTCAACGAAACCT

### **2.3.2 Identification des peptides provenant du gène de UBBP4**

L'identification des peptides uniques provenant du gène de UBBP4 (des deux cadres de lectures différents) a été faite en utilisant la base de données OpenProt. Rapidement, des expériences de spectrométrie de masse à large échelle provenant de différents tissus ont été réanalysées avec une base de données contenant les séquences des protéines de références et des protéines alternatives. Seulement les peptides uniques pour UBBP4 ont été considérés pour démontrer l'expression de la protéine.

### **2.3.3 Clonage et plasmides**

HA-Ubiquitine, HA-UBBP4-A1, A2, A3 et B1 ont été synthétisés par gblocks et inséré dans pcDNA 3.1 par digestion enzymatique XhoI-BamHI. Lamine A, Lamine B2 et PCNA ont été obtenus à partir d'une librairie d'ADNc provenant de cellules HeLa générée par RT-PCR avec un oligo-dT. Les ARNm avaient été préalablement extraits en utilisant le réactif TRIzol® (Invitrogen Life Technologies) en suivant les étapes fournies par la compagnie. Les différentes séquences ont été obtenues avec des oligos contenant les séquences ATTB permettant la recombinaison dans le plasmide pDONR221 (Invitrogen Life Technologies, Burlington, ON, Canada) lors d'une réaction utilisant la BP clonase II (Invitrogen Life Technologies) et par la suite lors d'une seconde ronde de recombinaison utilisant la LR clonase II (Invitrogen Life Technologies) pour les insérer dans pDEST47 (Invitrogen Life Technologies) afin de produire des protéines de fusion avec une GFP en C-terminal.

### **2.3.4 Traitement au MG132**

Les cellules ont été transfectées pendant 24h avec le vecteur vide (pcDNA3.1) ou avec le plasmide contenant HA-Ub ou HA-Ub<sup>KEKS</sup>. Par la suite, les milieux ont été changés pour du nouveau contenant soit du DMSO, soit 5µM de MG132 (Sigma-Aldrich) pour 12h.

### 2.3.5 Co-immunoprécipitations pour PCNA, LMNA et LMMB2

À la suite de la transfection des différentes constructions dans les cellules HeLa, les cellules étaient décollées en grattant les pétris dans du PBS1X froid et lysées dans du tampon Hight-Salt (HS) (1% NP-40, 50 mM Tris pH7.5, 300 mM NaCl, 150 mM KCl, 5 mM EDTA, 1 mM DTT, 10% glycerol, *Roche Complete Protease Inhibitor Cocktail*) sur glace pendant 10 minutes. Les extraits étaient par la suite soniqués 3 fois 10 secondes à 20% d'intensité avec un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific) et laissés un autre 10 minutes sur glace. Par la suite, les lysats étaient clarifiés par centrifugation à 13 000g à 4°C et à la suite du dosage des protéines, la même quantité étaient incubés avec des billes GFP (GFP-Trap-A, ChromaTek) pour 2h à 4°C. Après le temps d'incubation, les billes étaient lavées 3 fois avec du tampon HS et diluées dans du Laemli 1X.

### 2.3.6 CRISPR-Cas9 UBBP4 Knockout (KO)

Les cellules HeLa et U2OS délétées pour le gène UBBP4 ont été générées selon le protocole décrit dans (Ran et al. 2013) avec quelques modifications mineures. Plus précisément les guides d'ARN (ARNgs) ont été générés en utilisant l'outil en ligne du *The Broad Institute* (CIRSPRko). Les meilleurs candidats ont par la suite été confirmés avec un autre outil en ligne du *Centre for Organismal Studies* (CCTOP) et nous avons vérifié que ceux-ci ne ciblaient pas les gènes *UBB*, *UBC*, *UBA52* et *RPS27* ainsi que les autres pseudogènes de l'ubiquitine. Les oligos ont été hybridés et clonés dans le vecteur pSpCas9(BB)-2A-GFP (Addgene #48138, Cambridge, Massachusetts). Des paires d'ARNgs ont été utilisées afin d'enlever le gène UBBP4 au complet. Les plasmides validés par séquençage ont été transfectés dans des cellules HeLa et U2OS qui ont été sélectionnées après 24h par FACS (GFP) pour faire une sélection clonale. Les clones ont été vérifiés par PCR et ceux présentant un profil de délétion ont été confirmés par séquençage pour la région éditée.

### **2.3.7 Essai de prolifération par Carboxyfluorescein succinimidyl ester (CFSE)**

Pour les essais de prolifération, 125 000 cellules contrôles et UBBP4<sup>-/-</sup> ont été utilisées pour chaque jour de décompte. Les cellules ont été comptées et mises en suspension dans 1ml de PBS 1X et 1ml de PBS 1X contenant 10 $\mu$ M de CFSE (Biolegend) et incubées 20 minutes à 37°C. L'incorporation a été arrêtée en ajoutant 10ml de milieu DMEM complet et après centrifugation, le surnageant a été enlevé et remplacé par du milieu frais pour 10 minutes et incubé à la température de la pièce avant d'être placé dans des pétris de 60mm. Pour chaque temps (24, 48, 72 et 96h), les cellules ont été décollées en utilisant 500ul de trypsine et lavées 2 fois avec du PBS 1X froid avant d'être fixées dans 100% éthanol et conservées à -20°C. La quantification du CFSE a été faite par FACs (BD Fortessa cytometer, BectonDickinson) et les données ont été analysées par le logiciel FlowJo.

### **2.3.8 Quantification des nucléoles**

Pour la quantification de l'aire des nucléoles, le logiciel CellProfiler a été utilisé (<https://cellprofiler.org/>). Les noyaux ont été identifiés via un module d'identification des objets primaires en utilisant le DAPI comme marquage et les nucléoles via le marquage de la nucleolin. Chacun des nucléoles a été rapporté à son noyau via le module de connexion des objets et l'aire de chacune des structures a été déterminé en utilisant le module de mesure de la taille des objets.

### **2.3.9 Immunoprécipitation des protéines modifiées par HA-Ub et HA-Ub<sup>KEKS</sup>**

Pour les IP dans les conditions non dénaturantes, un pétri de 100mm des lignées stables a été induit pendant 48h avant que celles-ci ne soient lavées au PBS 1X et décollées avec la trypsine. Après un lavage au PBS1X, les cellules ont été lysées avec du HS (1% NP-40, 50 mM Tris pH7.5, 300 mM NaCl, 150 mM KCl, 5 mM EDTA, 1 mM DTT, 10% glycerol, *Roche Complete Protease Inhibitor Cocktail*) et laissées 10 minutes sur glace avant d'être soniquées 3 fois 10 secondes à 20% d'intensité

avec un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific) et laissées un autre 10 minutes sur glace. Les lysats ont par la suite été clarifiés par centrifugation et dosés afin qu'une quantité égale de protéines soit incubées avec l'hybridome HA (12CA5) pour 16h à 4°C. Des billes de protéine A sépharose ont été ajoutées et laissées en contact avec l'extrait pour 2h à 4°C. Les billes ont par la suite été centrifugées et lavées 3 fois au tampon HS.

Pour les IP dans des conditions dénaturantes, un pétris de 100mm de cellules qui expriment de manière stable les protéines de fusions ont été induites pendant 48h et lavées au PBS 1X avant d'être décollées à la trypsine. Les cellules ont par la suite été lysées avec du tampon A (2% SDS, 150mM NaCl, 10mM Tris-HCL pH 8.0 et un mélange d'inhibiteurs de protéases *Roche Complete Protease Inhibitor Cocktail*), bouillies à 95°C pour 5 minutes et soniquées 3 fois 10 secondes à 20% d'intensité avec un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific). Les échantillons ont par la suite été dilués avec 9 volumes de tampon B (10mM tris-HCL pH 8.0, 150mM NaCl, 2mM EDTA et 1% triton) et dosés pour qu'une quantité égale de cellules soient en contact avec l'hybridome HA (12CA5) pour une durée de 16h. Par la suite, des billes de protéine A sépharose ont été ajoutées pour une durée de 2h à 4°C et lavées 3 fois avec le tampon B.

### **2.3.10 Enrichissement des peptides modifiés avec GlyGly**

Les cellules HeLa normales ainsi que les cellules UBBP4<sup>-/-</sup> (les deux clones) ont été cultivées dans du milieu SILAC léger (HeLa de type sauvage), moyen (clone 2.7) et lourd (clone 4.3) pour un minimum de 6 passages. Une quantité suffisante de pétris ont été cultivés pour obtenir une concentration d'environ 6 mg de chacune des conditions afin d'obtenir environ 20 mg de protéines totales après le mélange des trois extraits cellulaires. Pour l'enrichissement de la signature GlyGly sur les peptides dans les cellules UBBP4<sup>-/-</sup>, nous avons utilisé la trousse disponible *PTMScan® Pilot Ubiquitin Remnant Motif (K-ε-GG)* (Cell Signaling Technology) en suivant le protocole fourni avec celui-ci.

### **3. Résultats**

#### **3.1 Chapitre 1 : Étude des cibles du complexe DDB1-Cul4 par spectrométrie de masse via l'intermédiaire des protéines DCAFs**

Les protéines MCMs (Mini-Chromosome Maintenance) forment un complexe protéique qui est très bien connu pour son rôle d'hélicase dans la réplication de l'ADN. En effet, les protéines MCM2 à MCM7 forment un complexe hétérohexamérique qui permet l'ouverture de la double hélice pour donner une matrice simple brin. En plus de son rôle d'hélicase, plusieurs des aspects du complexe, tels que son chargement à la chromatine ou encore son activation, sont des processus hautement régulés puisque ces étapes servent de point de contrôle entre les différentes phases du cycle cellulaire (Ishimi 2018). Au cœur de cette régulation, la modulation des partenaires d'interaction du complexe joue un rôle important en plus des nombreuses modifications post-traductionnelles qui peuvent survenir sur chacune des sous-unités des MCMs (Z. Li et Xu 2019).

Des études effectuées au laboratoire ont démontré que le complexe MCM en plus de son rôle dans la réplication, pourrait aussi jouer un rôle dans la réponse aux dommages à l'ADN lors de l'induction de cassures double-brin (Drissi et al. 2015). Dans cette étude, nous avons démontré que lors de dommages à l'ADN, les MCMs présentent une distribution cellulaire différente avec une plus large fraction des protéines du complexe liés à la chromatine comparativement aux cellules normales. De plus, nous avons identifié dans cette étude de nombreuses modifications post-traductionnelles sur chacune des sous-unités du complexe, qui présentent soit une augmentation soit une diminution en présence de dommages. Parmi celles-ci, nous avons identifié de nombreux sites d'ubiquitylation qui augmentent en réponse aux dommages à l'ADN sur plusieurs des protéines MCM.

Dans une seconde étude publiée par le laboratoire, nous avons démontré que de nombreuses protéines présentent une modulation de leur interaction avec les MCMs en réponse aux dommages à l'ADN (Dubois et al. 2016). Parmi celles-ci, on retrouve un complexe de E3, le complexe CRL4, qui interagit avec les MCMs en condition basale, mais qui augmentent son interaction en réponse aux cassures double-brin. Ces résultats suggèrent donc que le complexe MCM pourrait être ubiquitinylé en réponse aux dommages via le complexe CRL4.

Le complexe CRL4 fait partie de la grande famille des Cullins et se compose des protéines Cullin4a ou Cullin4b en association avec la protéine DDB1. Comme décrit précédemment, les membres de la famille des Cullins utilisent des protéines adaptatrices pour la reconnaissance des substrats ce qui leur permet de cibler une multitude de protéines dans la cellule. Le complexe CRL4 utilise les protéines DCAFs, comme protéines de reconnaissance pour les divers substrats à ubiquitinyler par le complexe. Près d'une soixantaine de protéines DCAFs ont été initialement rapportées par différentes études pour appartenir à cette famille de protéines. Bien que certaines des protéines DCAFs aient des rôles bien définis au sein de la cellule, d'autres restent néanmoins moins bien caractérisées et ont été placées dans cette famille de protéines seulement par la présence d'un domaine consensus qui est responsable de la reconnaissance de la protéine DDB1. En contrepartie, certaines protéines, telles que les protéines MCM, sont des substrats potentiels du complexe, mais avec aucune DCAF connue pour leur reconnaissance.

Dans l'optique d'identifier si le complexe DDB1-Cul4-DCAF13 est responsable de l'ubiquitylation des protéines MCM en plus de mieux caractériser chacune des protéines de la famille des DCAFs, nous avons décidé d'étudier chacune des protéines de cette famille par différentes approches de spectrométrie de masse afin de construire un profil pour chacune d'entre elles.



### 3.1.1 La famille des protéines DCAFs

Comme mentionné, les protéines DCAFs ont été étudiées par deux groupes de chercheurs qui ont utilisé différentes approches pour identifier les membres de cette famille de protéines (Angers et al. 2006; He et al. 2006). Le tableau 5 contient les noms des gènes correspondant aux protéines DCAFs qui ont été trouvées dans ces deux études et qui ont été rapportées dans une revue de la littérature en 2007 par Lee et al. (J. Lee et Zhou 2007b). Le tableau contient également les autres noms potentiels de chacune de ces protéines ainsi que le nom utilisé par la banque de données UniProt qui est celle que nous avons utilisée pour toutes nos analyses de spectrométrie de masse.

**Tableau 5 :Tableau des protéines DCAFs**

Gene ID	Others ID	Uniprot ID	Gene ID	Others ID	Uniprot ID
APG16L	UNQ9393/PRO34307	ATG16L	IFRG15	TOR1AIP2,LULL1	TRO1AIP2
COP1	RFWD2,,RNF200	RFWD2	KATNB1	-	KATNB1
CRBN	AD-006	CRBN	mβTrCP	BTRCP,FBW1A,,FBXW1A	BTRCP
CSA	ERCC8,CKN1,	ERCC8	NLE1	HUSSY-07	NLE1
DACF10	WDR32	DACF10	NUP43	-	NUP43
DBB2	-	DBB2	PAFAH1B1	LIS1,MDCR,MDS,PAFAHA	PAFAH1B1
DCAF1	KIAA0800,RIP,VPRBP	VPRBP	PWP1	-	PWP1
DCAF11	WDR23,GL014,PRO2389	DCAF11	RBBP4	RBAP48	RBBP4
DCAF12	KIAA1892,TCC52,WDR40A	DCAF12	RBBP5	RBQ3	RBBP5
DCAF13	WDSOF1,HSPC064	DCAF13	RBBP7	RBAP46	RBBP7
DCAF14	PHIP,WDR11	PHIP	SMU1	-	SMU1
DCAF15	C19orf72	DCAF15	TLE1	-	TLE1
DCAF16	C4orf30	DCAF16	TLE2	-	TLE2
DCAF17	C2orf37	DCAF17	TLE3	KIAA1547	TLE3
DCAF19	BRWD1,C21orf107,WDR9	BRWD1	TRPC4AP	C20orf188,TRRP4AP	TRPC4AP
DCAF2	DTL,CDT2,CDW1,L2DTL,RAMP	DTL	WDR12	-	WDR12
DCAF3	AMBRA1,KIAA1736	AMBRA1	WDR26	CDW2,MIP2,PRO0852	WDR26
DCAF4	WDR21,,WDR21A	DCAF4	WDR39	CIAO1,CIA1	CIAO1
DCAF5	BCRG2,KIAA1824,WDR22	DCAF5	WDR5	BIG3	WDR5
DCAF6	IQWD1,MSTP055	DCAF6	WDR51B	POC1B	POC1B
DCAF7	HAN11,WDR68	DCAF7	WDR53	-	WDR53
DCAF8	H326,WDR42A	DCAF8	WDR57	SNRNP40,PRP8BP,SFP38	SNRNP40
DCAF9	WDTC1,KIAA1037	WDTC1	WDR59	KIAA1923,FP977	WDR59
DDA1	C19orf58,PCIA1	DDA1	WDR5B	-	WDR5B
DET1	-	DET1	WDR61	-	WDR61
EED	-	EED	WDR76	-	WDR76
FBXW5	FBW5,PP3971	FBXW5	WDR82	TMEM113,WDR82A,UNQ9342	WDR82
FBXW8	FBW6,FBW8,FBX29,FBXO29,FBXW6	FBXW8	WSB1	SWIP1	WSB1
GRWD1	GRWD,KIAA1942,WDR28	GRWD1	WSB2	-	WSB2
Gβ2	-	GNB2			

### **3.1.2 Profil des protéines DCAFs**

#### **3.1.2.1 Expression des ARNm des DCAFs**

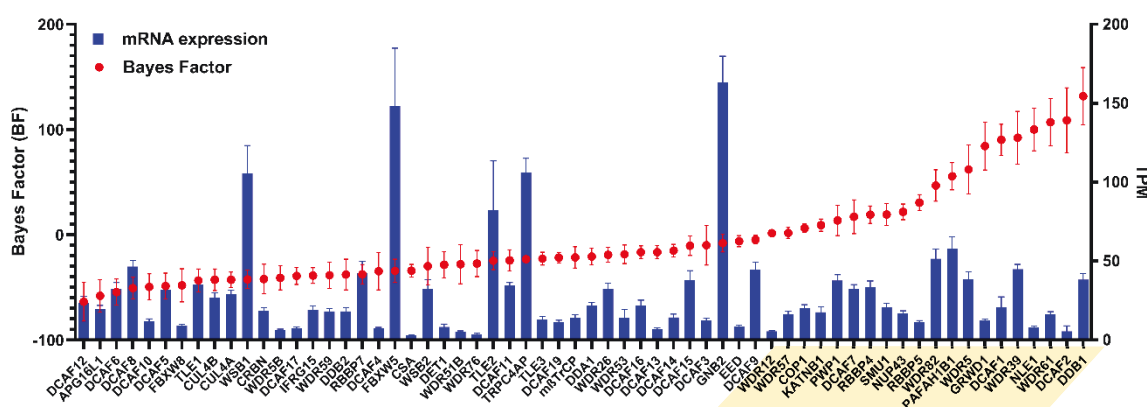
Les protéines DCAFs sont une famille de protéines très hétérogènes en ce qui concerne aussi bien leur localisation cellulaire que leurs fonctions. Ainsi avant même de commencer à caractériser leurs rôles cellulaires par différentes approches, nous avons décidé d'utiliser des données disponibles afin de construire un profil pour chacune d'entre elles. Les premières données que nous avons récupérées sont celles de l'expression des ARNm de tous les gènes des DCAFs dans différents tissus du corps humain. Pour ce faire, nous avons repris des données de séquençage d'ARNm disponibles via *Expression Atlas* pour chacune des DCAFs et nous avons aussi inclus les données pour les protéines DDB1, Cullin 4a et Cullin 4b (Petryszak et al. 2016). Sur la figure 11, la moyenne d'expression pour chacune des protéines DCAFs a été calculée et est représentée sur le graphique par les bars de couleurs bleues. Comme on peut le constater, il ne semble pas avoir de profil particulier en ce qui concerne l'expression de cette famille de protéines et c'est aussi le cas lorsqu'on analyse l'expression en fonction des différents tissus (résultats non montrés).

#### **3.1.2.2 Identification des protéines DCAFs essentielles à la survie cellulaire**

Puisque plusieurs de ces protéines ne sont pas encore très bien caractérisées, très peu de données sont disponibles quant à leurs fonctions dans la cellule. L'une des pistes qui peut nous renseigner sur l'importance de ces protéines est l'essentialité de celles-ci pour la survie cellulaire. Pour déterminer lesquelles de ces protéines sont essentielles à la cellule, nous avons repris une étude faite en 2015 par Hart et al. qui a évalué à large échelle quels gènes étaient essentielles ou non (Hart et al. 2015). Parmi les résultats qu'ils ont obtenus, nous avons recherché les données pour nos protéines d'intérêt et représenté la moyenne des Bayes factor (BF) pour chacune d'entre elles. Plus le BF est élevé, plus la protéine est essentielle pour la cellule. Afin de déterminer quelles protéines DCAFs sont essentielles, nous avons repris les mêmes critères qui ont été utilisés dans l'étude qui correspondent

à trois scores BF positifs parmi les cinq lignées cellulaires à l'étude. Nous avons aussi inclus les données pour la protéine DDB1 ainsi que pour les Cullins 4a et 4b. Nous avons classé les protéines dans un graphique en fonction de la moyenne de leur BF et nous avons surligné en jaune les protéines qui remplissent les critères de sélection.

De manière attendue, la protéine avec le BF le plus élevé est la protéine DDB1 avec une moyenne de 131,7 et un score positif dans les 5 lignées cellulaires, ce qui signifie que celle-ci est essentielle pour chacune d'entre elles (Figure 11). Parmi les autres protéines avec un score positif dans les 5 lignées cellulaires, on retrouve les protéines DCAF2, WDR61, NLE1, WDR39, DCAF1, GRWD1, WDR82 et PAFAH1B1. De plus, 10 autres DCAF répondent aux critères pour être classées comme gènes essentiels dans la cellule soit WDR5, RBBP5, NUP43, KATNB1, PWP1, WDR12, DCAF7, RBBP4, WDR57 et COP1. En tout, 18 protéines DCAFs (19 avec DDB1) remplissent les critères pour dire que celles-ci sont essentielles.



**Figure 11 : Profil d'expression et caractère essentiel des DCAFs**

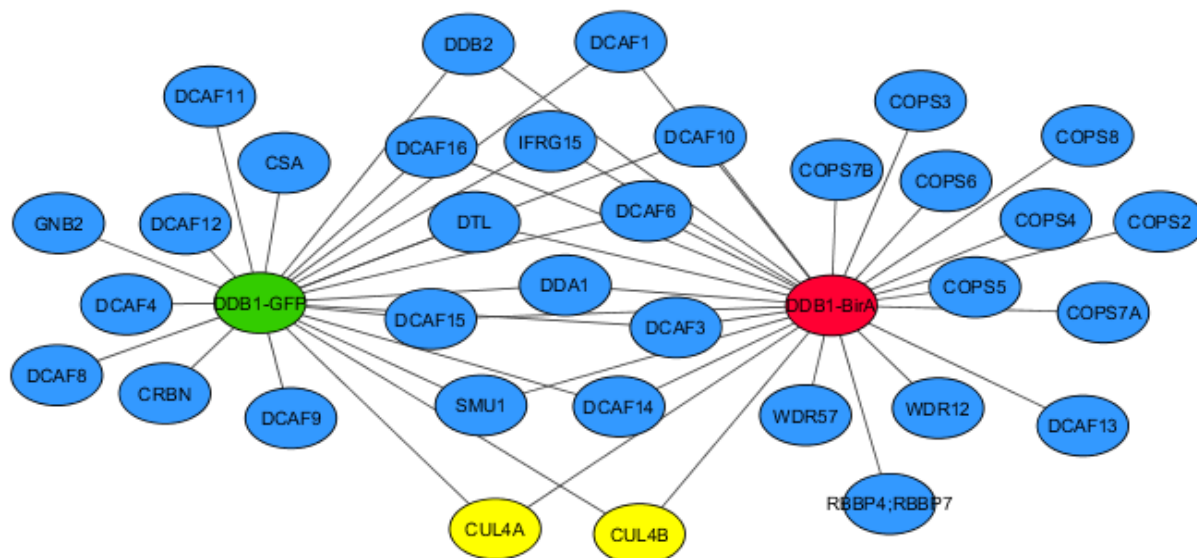
Les points rouges représentent la moyenne des Bayes Factor obtenus pour chacune des protéines DCAFs et la barre d'erreur représente la SEM (*standard error of the mean*). Pour que les protéines soient considérées comme essentielles, elles doivent avoir un BF supérieur à 0 dans 3 des 5 lignées cellulaires. Ainsi, celles qui remplissent ce critère ont été surlignées à l'aide d'un encadré jaune. Les barres en bleues représentent la moyenne en TPM (*transcripts per million*) de l'expression des ARNm pour chacune des protéines. La barre d'erreur représente elle aussi la SEM pour chacune des valeurs calculées.

### **3.1.3 Identification des protéines qui interagissent avec la protéine DDB1**

L'une des caractéristiques primordiales des protéines DCAFs est que celles-ci doivent interagir avec la protéine DDB1 pour former un complexe fonctionnel. Dans le but de faire une première identification des protéines qui appartiennent réellement à la famille des DCAFs, nous avons voulu identifier les protéines qui interagissent de manière constitutive avec la protéine DDB1. Pour y arriver, nous avons décidé d'enrichir la protéine DDB1 et d'analyser par spectrométrie de masse les partenaires d'interactions de celle-ci. Nous avons choisi deux méthodes d'enrichissement complémentaires pour identifier les protéines qui interagissent avec DDB1. La première méthode que nous avons utilisée est de fusionner la protéine DDB1 avec une GFP pour immunoprécipiter la protéine ainsi que ses partenaires d'interaction à l'aide de billes couplées à un anticorps reconnaissant l'étiquette. Cette approche permet d'identifier les protéines qui interagissent directement ou dans un environnement très rapproché de la protéine d'intérêt, et ce, au moment de l'immunoprécipitation. La deuxième méthode que nous avons utilisée est de fusionner DDB1 avec la protéine BirA\*. BirA\* est une biotine-ligase mutée qui permet d'attacher de manière covalente la biotine sur les protéines qui se retrouvent dans un rayon rapproché de la protéine. Ainsi, lorsque BirA\* EST fusionnée à DDB1, les protéines qui se retrouvent dans l'environnement de cette dernière sont biotinylées et donc en utilisant des billes de streptavidine, nous sommes en mesure de les identifier par spectrométrie de masse. L'avantage de cette approche, comparativement à la GFP, est que toutes les protéines qui sont dans l'environnement de la protéine d'intérêt pendant la période d'incubation avec la biotine peuvent être identifiées même si l'interaction est faible ou transitoire. Pour ces deux approches, nous avons décidé de générer des lignées stables dans les cellules U2OS en utilisant le système Flip-In. L'avantage principal de ce système est que celui-ci nous permet de connaître le site d'insertion du plasmide et nous assure donc aucun problème relié à l'intégration d'ADN exogène dans le génome. De plus, nous avons utilisé un plasmide qui permet de contrôler l'expression des protéines

de fusion via le répresseur à la doxycycline. Afin d'identifier les partenaires d'interactions de la protéine DDB1, nous avons utilisé le SILAC pour obtenir des ratios d'enrichissement pour toutes les protéines identifiées dans l'immunoprécipitation comparativement au contrôle.

Dans les immunoprécipitations avec la protéine de fusion GFP, nous avons obtenu un ratio d'enrichissement de 5,6 pour la protéine DDB1 et de 4,5 pour les protéines Cul4a et Cul4b. Pour ce qui est de l'approche avec la protéine de fusion avec BirA\*, les ratios pour DDB1, Cul4a et Cul4b sont de 8,23, 4,6 et 6,9 respectivement. Ainsi, nos deux méthodes ont été en mesure d'enrichir non seulement notre protéine d'intérêt, mais aussi les deux Cullins qui forment le complexe CRL4 avec elle. De plus, l'utilisation de la méthode BirA\* a permis d'identifier plusieurs protéines du complexe Cop9 signalosome qui sont importantes dans la régulation de l'activité du complexe (Figure 13). Pour ce qui est de l'identification des protéines DCAFs, la méthode GFP nous a permis d'identifier 20 protéines répertoriées dans la liste initiale alors que l'approche avec la protéine de fusion BirA\* nous a permis d'identifier seulement 15 d'entre elles (Fig.13). Plus précisément, huit de ces protéines ont été uniquement enrichies avec la fusion DDB1-GFP, 4 spécifiquement avec DDB1-BirA\* et 12 ont été identifiées par les deux approches.



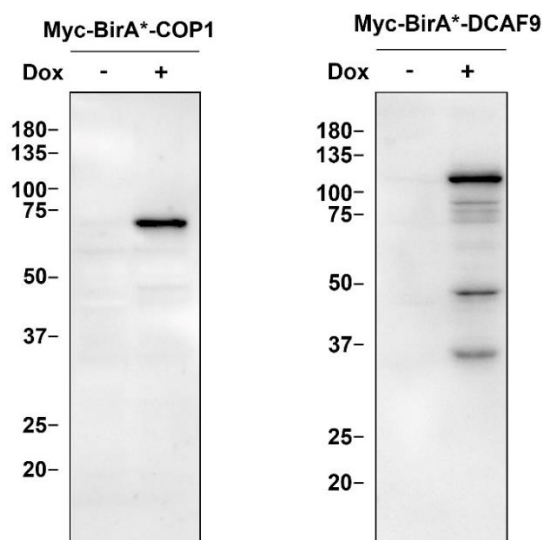
**Figure 12 : Identification des protéines DCAFs qui interagissent avec DDB1**

Des lignées cellulaires stables exprimant DDB1-GFP et DDB1-BirA\* ont été utilisées pour identifier, par spectrométrie de masse, les protéines DCAFs qui interagissent avec cette dernière. Les protéines qui sont reliées seulement avec la protéine DDB1-GFP ou DDB1-BirA\* ont été uniquement vues avec cette approche alors que celles qui se retrouvent au milieu ont été vues avec les deux méthodes. En jaune on retrouve les protéines Cul4a et Cul4B qui ont été identifiées avec les deux techniques.

### 3.1.4 Génération des lignées stables

Afin d'étudier plus en détail les protéines DCAFs, nous avons décidé d'identifier les partenaires d'interaction de chacune de celles-ci par spectrométrie de masse. Pour ce faire, nous avons dans un premier temps cloné toutes les protéines DCAFs dans un vecteur d'expression qui permet la production d'une protéine de fusion avec BirA\*. De plus, le plasmide contient un répresseur à la doxycycline qui nous permet de contrôler l'expression des protéines. À la suite de la confirmation par séquençage de chacun des plasmides, nous avons produit des lignées cellulaires stables dans des cellules U2OS en utilisant encore une fois le système Flip-In. La figure 14 montre quelques exemples de lignées cellulaires qui ont été générés par cette approche. De la doxycycline a été ajoutée aux cellules pendant 48h et les extraits cellulaires totaux ont été migrés sur gel SDS et transférés sur une membrane

de nitrocellulose. Les membranes ont par la suite été incubées avec un anticorps dirigé contre l'étiquette BirA\*. Tel qu'on peut constater sur la figure 13, à la suite de l'ajout de la doxycycline dans le milieu de culture, il y a apparition de bandes qui correspondent aux protéines d'intérêt avec un poids moléculaire correspondant à celui de la protéine plus celui de BirA\*.



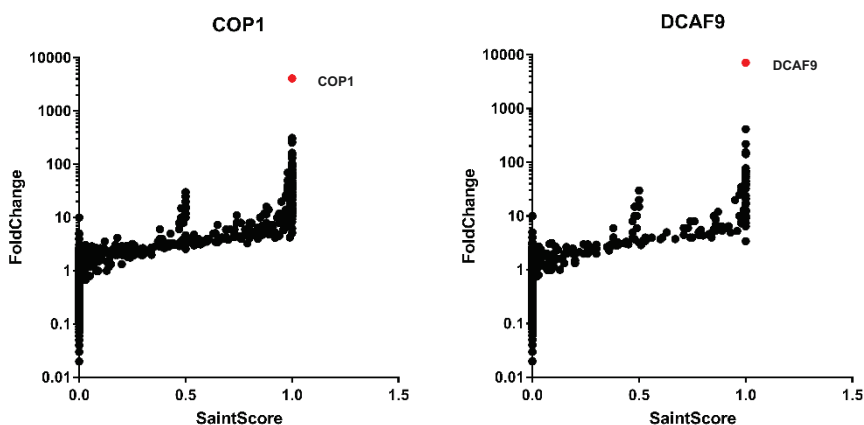
**Figure 13 : Preuve d'expression des protéines DCAFs-BirA\***

Extraits totaux de cellules U2OS sans ou avec ajout de doxycycline qui expriment différentes DCAFs fusionnées avec BirA\*. Les extraits ont été séparés sur un gel de SDS, transférés sur membrane avant d'être révélés avec un anticorps dirigé contre BirA\*.

### 3.1.5 Purification par affinité de l'interactome des DCAF

L'une des approches les plus utilisées afin de déterminer la fonction des protéines lorsque celle-ci est inconnue est la caractérisation de ses partenaires d'interaction par spectrométrie de masse. Nous avons donc décidé d'utiliser cette approche afin d'identifier un rôle potentiel pour chacune des protéines à l'étude. Pour y parvenir, nous avons induit l'expression de la protéine pendant 24h avant d'ajouter la biotine pour un autre 24h. Les protéines biotinylées ont par la suite été enrichies en utilisant des billes de streptavidine sépharose et les résultats ont été analysés en spectrométrie de masse. Des cellules U2OS vides ou des cellules

exprimant la protéine BirA\* seule ont été utilisées comme contrôle négatif pour identifier les interactions non spécifiques de cette méthode. Les résultats de chacun des réplicas ainsi que des contrôles ont été analysés par le logiciel SAINT pour déterminer quelles protéines sont enrichies de manière significative (Choi et al. 2011). Le logiciel SAINT permet de générer un score pour chacune des interactions allant de 0 (pas d'interaction) à 1 (interaction véridique) et calcule un enrichissement pour chacune des protéines (*FoldChange*). Des exemples des résultats obtenus par ce logiciel sont présentés à la figure 14. On remarque que la protéine qui présente le plus grand enrichissement (*FoldChange*) avec un score de 1 est la protéine d'intérêt qui est fusionnée à BirA\*. Ceci démontre que les expériences ont bel et bien marché et que notre méthode d'analyse est efficace pour identifier les protéines enrichies de manière significative. On remarque aussi un second enrichissement avec un score SAINT de 0.5 présent dans chacun des graphiques. Ces protéines correspondent à celles qui ont été enrichies dans seulement un des deux réplicas.



**Figure 14 : Identification des protéines qui interagissent de manière significative avec les protéines DCAFs.**

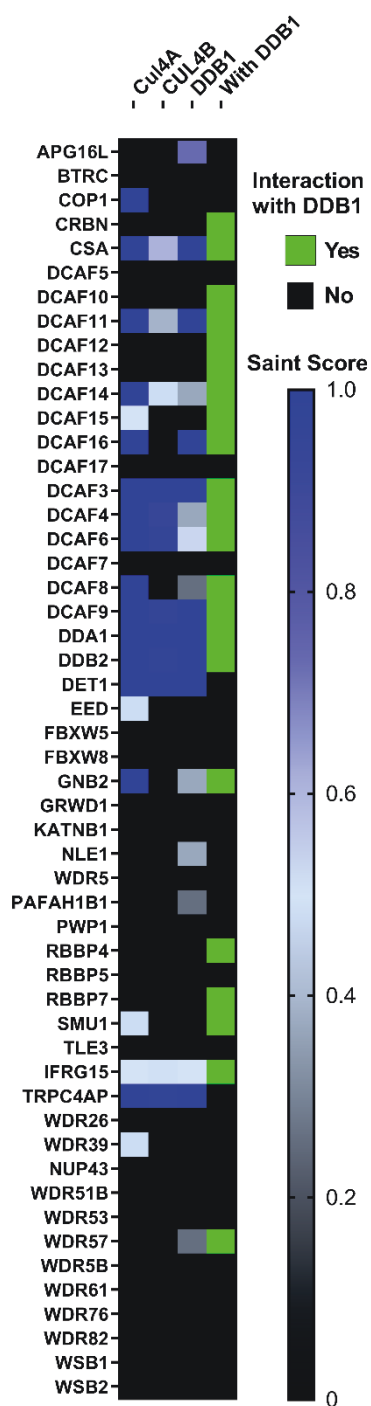
Graphique des scores obtenus grâce au logiciel SAINT en fonction du Fold Change pour chacune des protéines identifiées par spectrométrie de masse.



### **3.1.6 Interaction entre les DCAFs et les protéines du complexe CRL4**

Le premier but de cette étude est d'identifier les protéines qui peuvent servir de protéines adaptatrices pour le complexe DDB1-Cul4. Puisque dans un premier temps nous avons identifié les protéines qui interagissent avec la protéine DDB1 (IPs DDB1), nous nous sommes servis des résultats de chacune des purifications des DCAFs afin de déterminer si à l'inverse, les protéines du complexe CRL4 interagissaient avec chacune des DCAFs. Pour ce faire, nous avons pris les scores SAINT obtenus pour DDB1 et les protéines Cul4A et Cul4B dans chacune des purifications par affinité. Le sommaire des résultats est représenté dans la figure 15 sous la forme d'un Heatmap avec la couleur noire qui signifie l'absence d'interaction, jusqu'à une coloration bleu foncé qui correspond à une interaction avec un score de 1. Nous avons aussi intégré les données obtenues pour DDB1 (IP avec GFP et les purifications avec la streptavidine) dans le graphique. Puisque ces résultats sont exprimés avec des ratios, seules la présence ou l'absence d'interaction ont été rapportées (noir signifie aucune interaction et vert indique une interaction).

On constate en regardant la figure 15 la présence de certaines des protéines DCAFs qui ont été identifiées avec la protéine DDB1. Ceci confirme la réciprocité de l'interaction puisqu'on retrouve les composantes du complexe CRL4 avec la protéine DCAF. D'autres DCAFs ont cependant été identifiés avec DDB1, mais aucun des membres du complexe n'a été retrouvé lors de l'étude des leurs partenaires d'interaction. Quelques DCAFs ont identifié une ou plusieurs des composantes du complexe CRL4 alors que celles-ci n'avaient pas été identifiées lors de la première analyse avec DDB1. Finalement, une proportion des protéines DCAFs n'ont été vues ni avec la protéine DDB1 ni lors de l'étude des partenaires d'interaction de chacune d'entre elles.

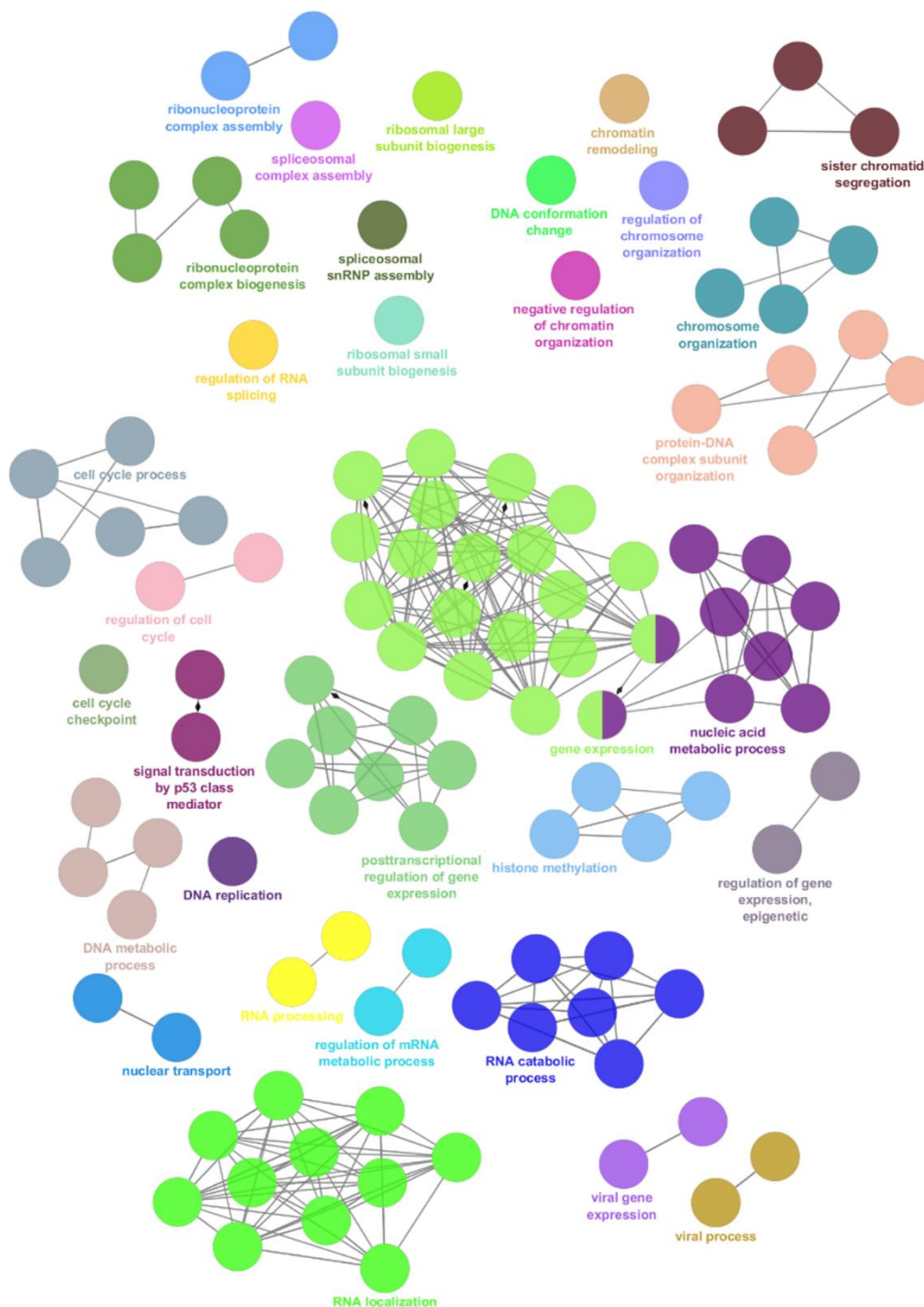


**Figure 15 : Heatmap des interactions entre les DCAFs et les composants du complexe CRL4**

Score SAINT pour les composants DDB1, Cul4a et Cul4b pour chacune des protéines DCAFs qui a été étudiées par spectrométrie de masse. La dernière colonne du graphique correspond aux résultats obtenus avec la protéine DDB1 fusionnée avec GFP et BirA\*.

### 3.1.7 Fonctions associées aux protéines DCAFs

À la suite de l'identification des protéines DCAFs qui interagissent avec le complexe DDB1-Cul4, nous avons voulu déterminer les fonctions cellulaires du complexe de E3 ubiquitine-ligase. Afin d'obtenir un portrait global de la fonction du complexe DDB1-Cul4, nous avons sélectionné les protéines qui présentent un enrichissement significatif dans chacune des IP des DCAFs ( $FDR < 0,05$ ) et nous avons enrichi les fonctions cellulaires à l'aide d'un logiciel de *Gene Ontology ClueGo* (Bindea et al. 2009). Puisqu'un certain nombre de protéines (1595 exactement) répondaient aux critères, nous avons sélectionné les fonctions cellulaires qui présentaient un enrichissement significatif ( $p\text{-value} < 1 \times 10^{-10}$ ) avec un minimum de 10 gènes qui représentent un minimum de 15% de l'ensemble des gènes totaux de la fonction. Avec ces critères, nous avons identifié 33 processus cellulaires qui sont représentés par les différentes couleurs à la figure 16. Les fonctions qui sont similaires ont été regroupées ensemble et seulement le nom du processus qui était le plus significativement enrichi a été affiché pour alléger la figure. Ainsi, plus il y a de cercles de la même couleur, plus les protéines associées à cette fonction ont été retrouvées enrichies dans des fonctions similaires à celle indiquée.



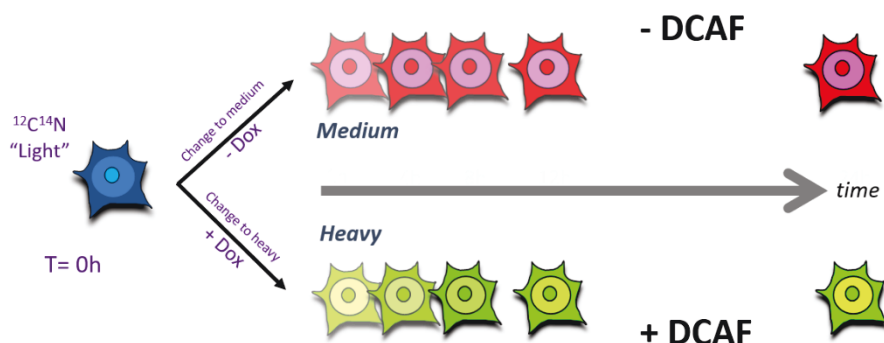
**Figure 16: Fonctions cellulaires associées aux complexes DDB1-Cul4-DCAFs**

Fonctions des protéines DCAFs qui ont été identifiées pour être associées aux complexes DDB1-Cul4. Les fonctions présentées ont toutes un p-value inférieur à  $1 \times 10^{-10}$ , un minimum de 10 gènes associés à la fonction et qui représentent un minimum de 15% des gènes totaux du processus.

### 3.1.8 Pulse SILAC

La majorité des événements d'ubiquitylation qui ont lieu dans la cellule mènent à la disparition de la protéine cible via la dégradation par le protéasome. Ainsi, la surexpression des composantes d'un complexe d'ubiquitine-ligase peut mener à la disparition du substrat et donc empêcher ou du moins limiter grandement sa détection en spectrométrie de masse. Afin de pallier à ce problème, une équipe de chercheur a développé une méthode de Pulse-SILAC afin de mesurer la stabilité des protéines dans des extraits totaux lors de traitements (An et al. 2017). Leur technique consiste à comparer les ratios SILAC (M/L) obtenus pour les protéines dans des conditions normales et de les comparer avec les ratios SILAC (H/L) dans des cellules qui ont été traitées avec différentes drogues. En comparant les deux ratios, on peut identifier les variations dans le niveau d'abondance des protéines. Nous avons décidé d'utiliser cette méthode pour identifier les changements d'abondance des protéines lors de la surexpression des protéines DCAFs afin de tenter d'identifier les substrats de celles-ci.

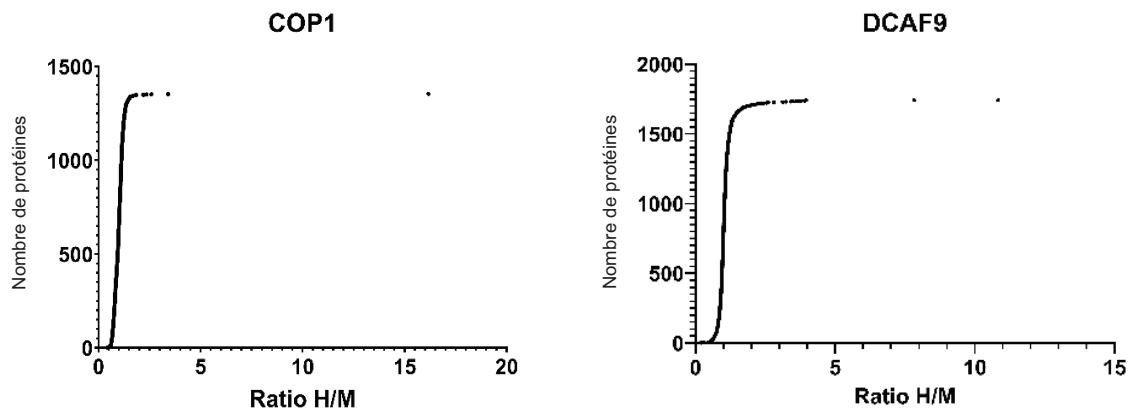
Nous avons repris les lignées cellulaires stables qui expriment chacune des protéines DCAFs et nous les avons cultivées dans du milieu SILAC léger (R0K0) dans deux différents pétris. Au temps 0, nous avons ajouté de la doxycycline à l'un des deux pétris et après 8h, nous avons changé le milieu du pétri contrôle pour du milieu moyen (R6K4) et celui du pétri qui exprime la DCAF avec du milieu lourd (R10K8). Après 24h total nous avons lysé les cellules, mélangé les extraits avec un ratio 1:1 et analysés les échantillons par spectrométrie de masse (Figure 17).



**Figure 17 : Schéma de l'approche Pulse-SILAC**

Les cellules ont été initialement cultivées dans du milieu léger soit du R0K0. Au temps = 0h, la doxycycline a été ajoutée aux cellules pour permettre l'induction de la protéine DCAF. Par la suite, le milieu de culture de ces cellules a été changé au temps = 8h pour du milieu moyen dans les cellules contrôles et dans du milieu lourd pour les cellules qui expriment la protéine DCAF. Par la suite les cellules ont été lysées, mélangées dans les mêmes proportions et à la suite de leur préparation, les échantillons ont été analysés en spectrométrie de masse.

Pour identifier les protéines qui sont diminuées lors de la surexpression d'une DCAF, nous avons considéré les protéines identifiées dans les deux répliques et donc le ratio H/M augmente. Ainsi, nous avons calculé un ratio moyen H/M pour chacune des protéines que nous avons classifié par la suite du plus petit au plus grand. La figure 18 présente des exemples de résultats que nous avons obtenus pour chacune des protéines DCAF. Comme on peut constater, la majorité des protéines possèdent un ratio de 1 ce qui signifie que la plupart ne voient pas leur abondance changer avec la production de notre protéine d'intérêt. Nous avons identifié seulement quelques protéines pour chacune des expériences qui ont des ratios H/M supérieurs à une valeur de 1,5 et qui ont donc une abondance moduler à la suite de l'expression de la DCAF.

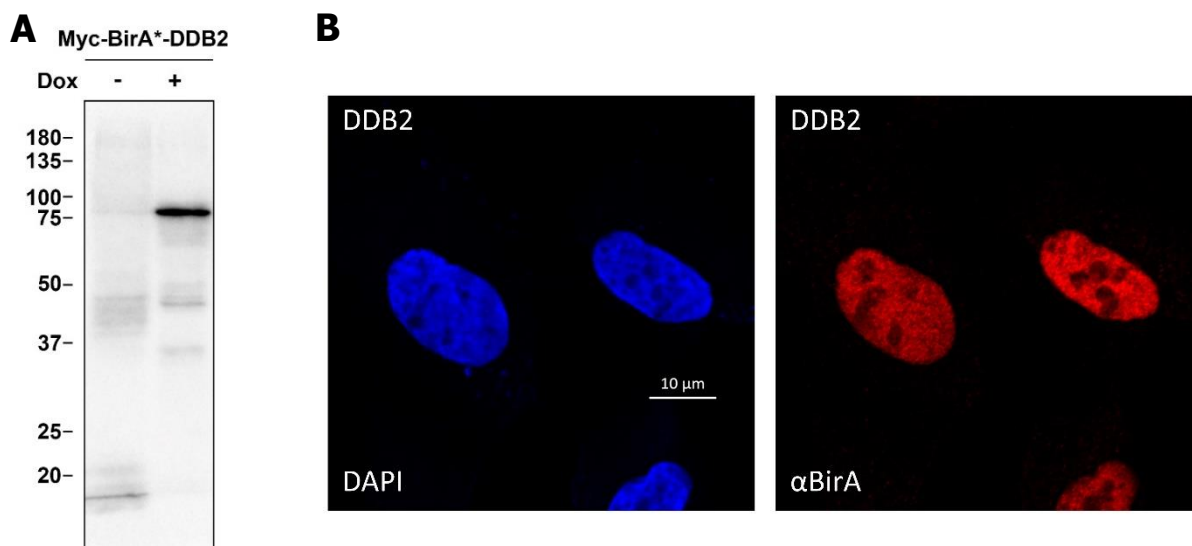


**Figure 18 : Pulse SILAC des protéines DCAFs**

Les graphiques ci-dessus représentent ces ratios en fonction de leur rang de grandeur. On peut voir par exemple pour la protéine COP1 que près de 1400 protéines ont été identifiées dont la plupart avec un ratio moyen d'environ 1. Pour la protéine DCAF9 on peut voir que plus de 1600 protéines ont été identifiées dans les deux répliques avec toujours une majorité avec un ratio moyen de 1.

### 3.1.9 Exemple de DDB2

La protéine DDB2 est un bon exemple des résultats que nous avons obtenus avec cette étude. Si l'on reprend plus en détails les résultats obtenus pour cette DCAF particulière, l'immunobuvardage des lignées cellulaires stables exprimant DDB2-BirA\* démontre la présence d'une bande avec un poids moléculaire légèrement supérieur à 75kDa lors de l'ajout de la doxycycline. Cette bande correspond au poids de la protéine DDB2 (47kDa) fusionné à la protéine BirA\* (37kDa). Afin de nous assurer de la bonne localisation de la protéine de fusion, nous avons décidé d'effectuer une immunofluorescence contre l'étiquette BirA\*. Tel que le démontre la figure 19, la protéine DDB2-BirA\* est strictement localisée au niveau du noyau des cellules, ce qui est aussi le cas pour la protéine DDB2 endogène. Ces résultats nous indiquent donc que la protéine de fusion est exprimée et qu'elle se localise au bon endroit dans la cellule .

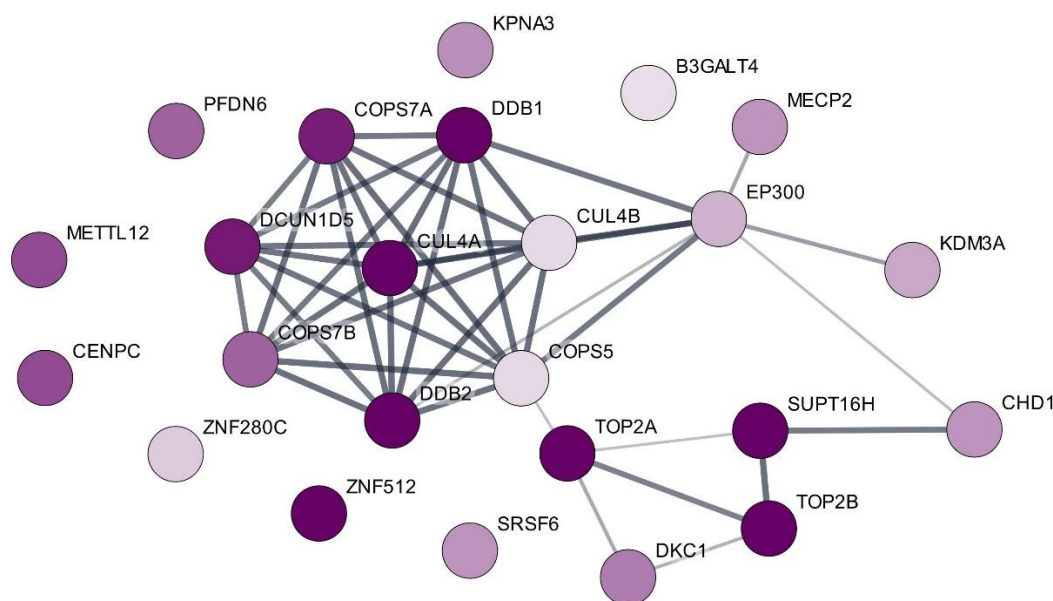


**Figure 19 : Confirmation des lignées cellulaires DDB2-BirA\***

A) Les lignées cellulaires U2OS DDB2-BirA\* ont été induites ou non et les extraits cellulaires ont été lysés dans du Laemli 1X directement et migrées sur gel SDS avant d'être transférées sur une membrane de nitrocellulose et hybridées avec un anticorps dirigé contre la protéine BirA\*. B) Des immunofluorescences ont été effectuées sur les cellules DDB2-BirA\* avec un anticorps dirigé contre l'étiquette. En rouge, on peut voir la localisation de la protéine de fusion qui est strictement nucléaire lorsqu'on compare avec le marqueur de noyau le DAPI en bleue.

Lors de l'identification des partenaires d'interaction de DDB2 par spectrométrie de masse, 24 protéines ont été identifiées avec un score SAINT de plus de 0.9 et avec un *False-Discovery-rate* inférieur à 0.01. En utilisant la base de données STRING, les interactions déjà répertoriées entre chacune de ces protéines ont été mises en évidence dans la figure 17. On constate que parmi ces protéines on retrouve les composantes du complexe CRL4 dont les protéines DDB1, Cullin4A et Cullin4B. De plus, d'autres protéines relatives aux fonctions des complexes d'ubiquitine ligase de la famille des Cullins, telles que les protéines COPS7B, COPS5 et COPS7A qui appartiennent au complexe COP9 signalosome, ont été identifiées.



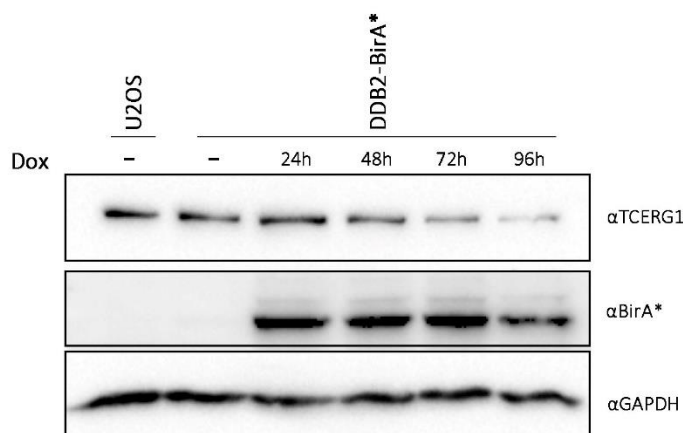


**Figure 20 : Interactome de la protéine DDB2-BirA\***

Interactome des protéines qui ont eu un Score SAINT supérieur à 0.9 avec la protéine DDB2-BirA\* obtenu grâce au logiciel Cytoscape. Les liens représentent les interactions déjà répertoriées dans la base de données STRING, plus il y a d'évidence pour l'interaction dans la littérature, plus le lien va être de couleur foncé et épais. De plus, la coloration de chacun des cercles représente le FoldChange calculé pour chacune des protéines. Plus la coloration est foncée, plus le FoldChange est élevé et donc plus la protéine a été enrichie dans les purifications par affinité.

Pour les expériences de Pulse-SILAC de la protéine DDB2, 1083 protéines ont été identifiées dans les deux conditions (M/L et H/L) ainsi que dans chacun des réplicas. Parmi celles-ci, 29 présentent un ratio moyen H/M supérieur à 1,7 et 28 d'entre elles ont été identifiées avec plus de 2 peptides uniques. Dans cette liste, on retrouve la protéine TCERG1 qui est une protéine impliquée dans la transcription et dans l'épissage alternatif. Pour confirmer les résultats obtenus en spectrométrie de masse, nous avons surexprimé la protéine DDB2-BirA\* à différents temps avec l'aide de la doxycycline (0,24,48,72 et 96h) et observé les niveaux totaux de la protéine TCERG1. Comme le présente la Figure 21, on peut clairement constater que le niveau de la protéine TCERG1 diminue de plus en plus en fonction du temps. Ceci nous confirme que l'expression de la protéine DDB2 amène une diminution de la

quantité de la protéine TCERG1 et renforce les résultats obtenus à l'aide de l'approche de Pulse-SILAC.



### Figure 21 : Schéma de l'approche Pulse-SILAC

Les lignées DDB2-BirA\* ont été induites à la doxycycline pour une période de temps allant de 24 à 96h. Les cellules ont par la suite été lysées dans du Laemli 1x et migrées sur gel SDS et transférées sur une membrane de nitrocellulose. Les membranes ont été incubées avec un anticorps BriA\* (pour confirmer l'expression de la protéine DDB2), un anticorps TCERG1 et finalement avec un anticorps GAPDH pour s'assurer du niveau des extraits cellulaires.

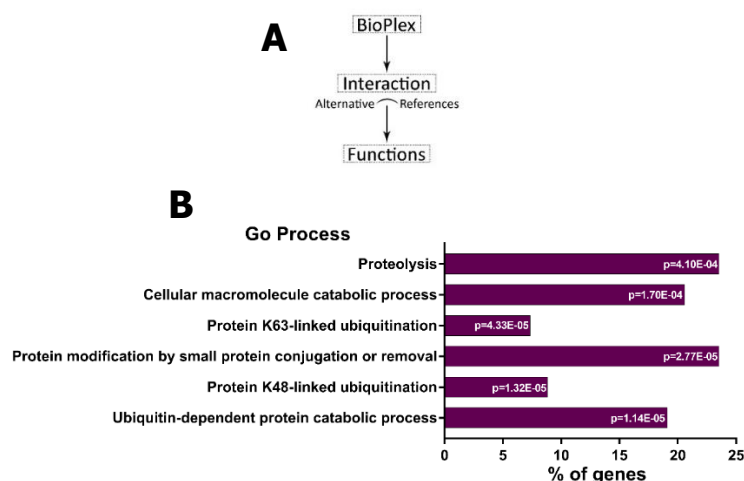
### **3.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène**

À la suite du séquençage du génome humain et à la quantité d'information générée par cette avancée scientifique, les chercheurs ont dû établir des critères stricts afin d'annoter le génome de la manière la plus rigoureuse possible. Ainsi, pour qu'une séquence d'acide nucléique soit annotée comme un gène, ils se sont basés sur trois critères principaux; le premier critère était que la protéine produite à partir du gène devait posséder une longueur minimum de 100 codons. La seconde condition était que le gène devait commencer par un codon d'initiation classique, c'est-à-dire un ATG et finalement, un seul cadre de lecture était annoté par gène et était impérativement celui qui produisait la plus longue protéine. Ainsi, suivant ces règles, le génome a été annoté pour déterminer ce qui était codant ou non.

Toutefois, de plus en plus d'études sont publiées démontrant l'expression de nouveaux cadres de lecture (ORF) qui se situent dans des ARN non codants ou encore dans des séquences régulatrices des ARNm (Andreev et al. 2015; Ingolia 2016). De plus, non seulement leur expression a été prouvée par diverses approches, mais pour certain, une fonction cellulaire leur a déjà été attribuée (Pauli et al. 2014). Ainsi, les critères initialement utilisés pour déterminer le pouvoir codant d'une séquence étaient probablement trop rigides pour représenter la réalité du génome. Une autre preuve qui vient appuyer cette affirmation est la découverte faite dans le laboratoire du Pr Roucou qu'un transcrit peu avoir plusieurs cadres de lecture (Vanderperre et al. 2011). En fait, non seulement les ARNm peuvent avoir des ORFs dans leur séquence régulatrice, mais d'autres cadres de lecture peuvent se chevaucher au sein de la séquence codante. En compilant les évidences de spectrométrie de masse et de *ribosome profiling*, ils ont ainsi construit une base de données regroupant les évidences d'expression de ces protéines alternatives (altORF) (Brunet et al. 2019).

### 3.2.1 Découverte d'UBBP4-B1

L'un des défis dans la découverte d'un grand nombre de protéines est non seulement de prouver leur existence et de démontrer que celles-ci ont une fonction au sein de la cellule, mais aussi de choisir un candidat à étudier. Dans cette optique, les résultats d'une étude à large échelle d'interactions protéiques publiées par Huttlin et *al.* en 2015 (BioPlex) ont été réanalysés en utilisant la base de données contenant les séquences des protéines alternatives (Figure 22) (Huttlin et al. 2015a). La rationalité derrière cette approche était de trouver un groupe de protéines de références (les protéines annotées dans le génome) qui interagissent avec une protéine alternative et de ce fait, pouvoir trouver une fonction à la protéine alternative. En utilisant cette approche, nous avons été en mesure d'identifier une protéine alternative, nommée IP\_710742 qui présente plusieurs interactions avec des protéines de références. L'analyse des fonctions de ces protéines par *gene ontology*, a démontré que celles-ci sont majoritairement impliquées dans des mécanismes relatifs à l'ubiquitine figure 22. La protéine IP\_710742 provient du gène UBBP4 (pour *UBB-pseudogene 4*) qui est répertorié comme un pseudogène du gène UBB qui encode pour l'ubiquitine.

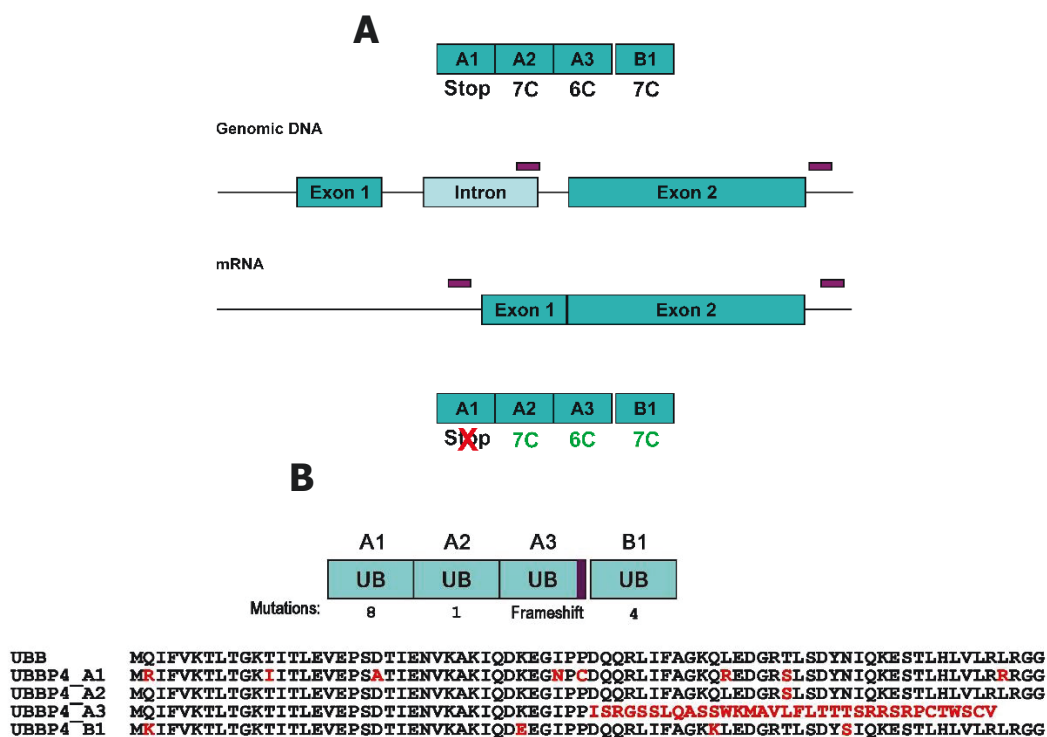


**Figure 22 : Découverte d'une protéine alternative IP\_710742**

A) Schéma de l'approche expérimentale utilisée pour trouver des groupes de protéines de références interagissant avec des protéines alternatives afin d'identifier une fonction cellulaire à celles-ci. B) Analyse de la fonction des protéines retrouvées pour interagir avec la protéine alternative IP\_710742

Le gène UBBP4 a initialement été annoté comme un pseudogène puisque malgré le fait que celui-ci encode pour trois répétitions de l'ubiquitine tout comme le gène parental (nommé A1 à A3, Figure 23), la présence d'une mutation dans la première ubiquitine amène l'apparition d'un codon-stop à la position 59. La terminaison prématurée de la protéine à cette position fait que les deux glycines en C-terminal ne sont pas présentes et que celle-ci ne peut donc pas s'attacher à une autre protéine (Cowland, Wiborg, et Vuust 1988). De plus, l'ubiquitine contient généralement une répétition de 7 cytosines dans sa séquence nucléotidique, la troisième répétition de l'ubiquitine chez UBBP4 n'en contient cependant que 6, ce qui amène un décalage du cadre de lecture et la production d'une protéine avec une séquence complètement différente. Ainsi, pour ces raisons (majoritairement dû à la présence du codon-stop), le gène UBBP4 a été annoté comme un pseudogène. Cependant, les bases de données disponibles ne présentent pas la même séquence pour le gène UBBP4 que celle qui a initialement été rapportée lors de sa découverte et celles-ci diffèrent entre autres, par l'absence du codon-stop dans la première répétition. Ainsi, pour déterminer quelle séquence d'acide nucléique était bel et bien présente au sein du génome, nous avons extrait l'ADN génomique ainsi que l'ARNm de UBBP4 provenant de 3 lignées cellulaires différentes (U2OS, 293 et HeLa) que nous avons fait séquencer. Il s'est avéré qu'une tyrosine était présente à la position 59 (tel que les bases de données rapportent) et non un codon-stop tel qu'initialement annoté (Figure 23). La cytosine est toujours manquante dans la 3<sup>e</sup> répétition ce qui mène quand même à la production d'une séquence qui ne ressemble pas à l'ubiquitine en C-terminal. Néanmoins, le gène UBBP4 peut produire 2 ubiquitines fonctionnelles dans son premier cadre de lecture et une troisième, la protéine alternative IP\_710742 (identifié UBBP4-B1), qui est présente dans un cadre de lecture +2 et est donc décalé de 1 acide nucléique comparativement au cadre de lecture annoté.

De manière intéressante, toutes les protéines produites par le gène comportent des mutations comparativement à la séquence canonique de l'ubiquitine (Figure 23). En fait, A1 comporte 8 acides aminés différents, A2 possède une seule mutation et on en dénombre 4 pour B1. Pour la répétition A3, bien que plus de la moitié de la protéine ait la même séquence, le manque de la cystéine telle que décrit précédemment entraîne un décalage du cadre de lecture, ce qui mène à la production d'une séquence en acides aminés qui n'est pas associée à l'ubiquitine.

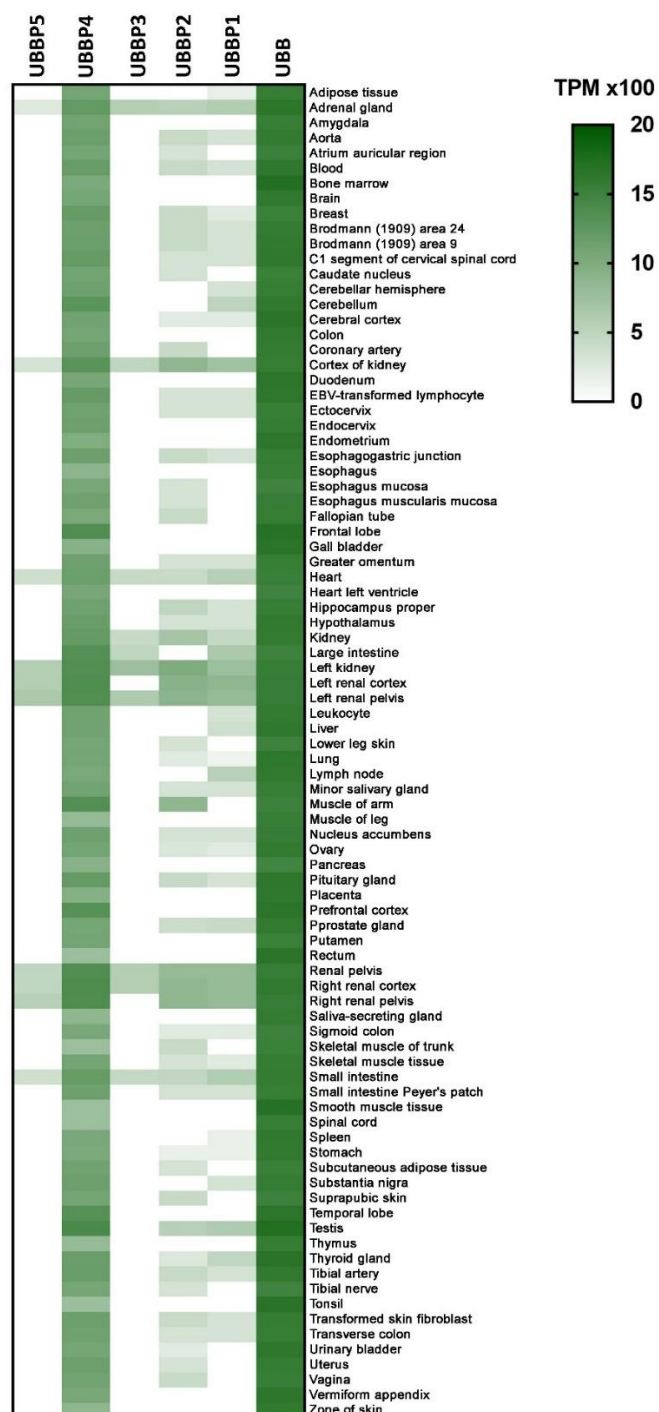


**Figure 23 : Le gène UBBP4 ne possède pas de codon-stop et n'est donc pas un pseudogène**

A) Schéma récapitulatif de l'approche utilisée pour le séquençage de l'ARNm et de l'ADN génomique pour UBBP4. Les résultats ont confirmé qu'UBBP4 ne possède pas un codon-stop à la position 59 comme il avait été rapporté dans la littérature, mais possède plutôt une tyrosine à cette position. De plus, les résultats de séquençage ont confirmé qu'il y avait seulement 6C dans la 3<sup>e</sup> répétition d'ubiquitine, ce qui mène à un décalage du cadre de lecture. B) Alignement de séquence de chacune des ubiquitines produites à partir du gène UBBP4 et de l'ubiquitine canonique. En rouge, les acides aminés qui diffèrent entre les séquences des ubiquitines de UBBP4 et de la canonique.

### 3.2.2 Expression de l'ARNm du gène UBBP4 dans les cellules humaines

Notre défi à la suite de cette découverte était de prouver qu'UBBP4 n'était pas un pseudogène et que celui-ci pouvait potentiellement produire des protéines fonctionnelles. La première approche que nous avons utilisée était de comparer l'expression de l'ARNm de ce gène comparativement aux autres membres de sa famille. En fait, UBBP4 provient d'une famille de pseudogènes composés de 5 membres nommés UBBP1 à UBBP5 qui dérivent tous du même gène parental UBB. Afin de comparer les expressions de chacun des membres de cette famille, nous avons repris des données de séquençage d'ARNm disponibles sur une plateforme en ligne (*Gtex bioportal*) et nous les avons comparés avec l'expression d'UBB. Pour être certain de pouvoir bien comparer les niveaux entre les différents ARNm, nous avons sélectionné les études qui prenaient en compte seulement les lectures de séquences uniques, c'est-à-dire où les séquences ne peuvent s'apparier que sur un seul endroit du génome. En effet, puisqu'il y a beaucoup de séquences en commun entre ces différents gènes, il est important de s'assurer que le signal obtenu provient bel et bien d'un seul gène. Basés sur ces critères, nous avons sélectionné 5 études à larges échelles qui ont évalué la quantité d'ARNm exprimés dans de nombreux types de tissus. Tel que le démontre la figure 24, les gènes UBBP1 et UBBP2 présentent un profil d'expression similaire avec une faible expression pour environ la moitié des tissus étudiés et une absence totale pour l'autre moitié. L'ARNm correspondant aux gènes UBBP3 et UBBP5 est absent de pratiquement tous les tissus analysés hormis pour quelques-uns qui présentent une très faible expression de ceux-ci. Comparativement aux autres membres de cette famille de pseudogène, l'ARNm de UBBP4 est plus fortement exprimé dans l'ensemble des tissus qui ont été étudiés. Son expression reste toutefois moindre que l'ARNm correspondant au gène parental UBB. Ainsi, ces résultats nous indiquent non seulement que l'ARNm de UBBP4 est présent dans de nombreux tissus humains et donc potentiellement traduit en protéine, mais que celui-ci est nettement plus exprimé que le niveau attendu pour un pseudogène sans séquence codante.



**Figure 24 : Expression des ARNm de la famille UBBP et d'UBB**

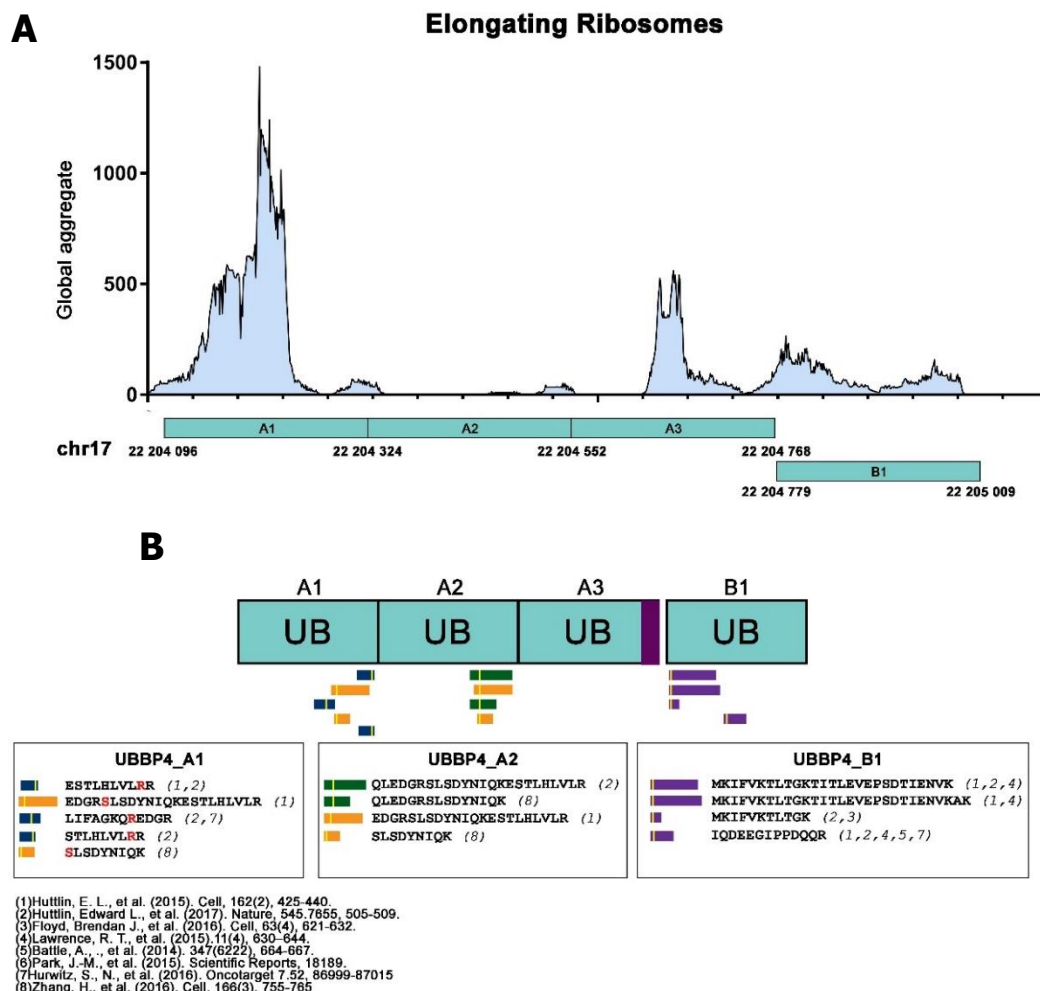
Analyse du niveau d'expression des ARNm du gène UBB et des pseudogènes UBBP1 à 5 provenant de différents tissus à l'aide de données de séquençage d'ARNm qui sont disponibles via le portail GTEx.



### 3.2.3 Traduction de l'ARNm d'UBBP4

L'étape suivante était de prouver la production des protéines à partir de l'ARNm d'UBBP4, et ce, pour les deux cadres de lectures. Dans un premier temps, nous avons voulu déterminer s'il y avait la présence ou non de ribosome sur son ARNm. Pour ce faire, nous avons récupérés des études de *ribosomes profiling* déjà publiées et nous avons extraits les données pour la section du génome qui encode pour UBBP4. Tel que démontré à la figure figure 25, nous avons un signal marqué pour le début du gène qui correspond à la répétition A1, une absence de signal pour la seconde partie et un regain d'intensité pour la partie C-terminal de A3. En ce qui concerne le deuxième cadre de lecture qui contient B1, on obtient un signal de plus faible intensité que pour A1, mais celui-ci est présent pour toute la longueur de la séquence qui code pour B1. Ainsi, ces résultats démontrent la présence de ribosome d'élongation sur toute la longueur de l'ARNm du gène UBBP4 y compris la portion contenant la répétition B1.

Par la suite, nous avons voulu confirmer la traduction en protéine de cet ARNm par la détection de peptides correspondant aux différentes protéines du gène UBBP4. Puisque chacune de ces protéines contient des mutations dans sa séquence d'acides aminés comparativement à l'ubiquitine canonique, il est possible d'identifier des peptides uniques pour chacune d'entre elles. Tout comme pour l'expression en ARNm, il est impératif de s'assurer que les peptides détectés ne peuvent s'apparier à aucune autre protéine dans la cellule. En utilisant la base de données *openprot*, nous avons facilement pu identifier les peptides uniques qui ont été répertoriés pour chacune des protéines d'UBBP4. La figure 25 schématise les différents peptides identifiés pour ces protéines avec leur emplacement approximatif sur le gène. On peut voir que pour la répétition A1, 5 peptides uniques ont été détectés et que 4 peptides ont été identifiés pour les répétitions A2 et B1 respectivement. Aucun peptide pour la protéine A3 n'a cependant été détecté dans ces analyses de spectrométrie de masse.



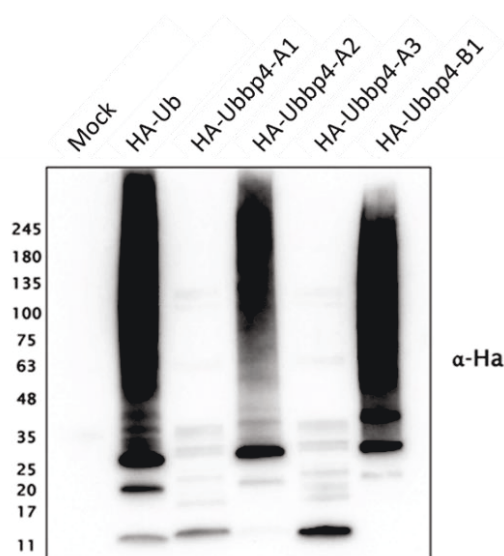
**Figure 25 : Preuve de l'expression de 4 différentes protéines à partir du gène UBBP4 dans deux cadres de lecture différents**

A) le ribosome profiling utilise une approche par séquençage pour identifier les fragments d'ARNm qui sont protégés lors d'une digestion enzymatique par les ribosomes d'élongation. Le schéma en A représente les signaux obtenus pour les fragments d'ARNm qui proviennent du gène UBBP4 à partir d'une base de données qui compile des résultats provenant de différentes études de ribosome profiling (GWIPS-viz). Les signaux obtenus couvrent non seulement le premier cadre de lecture qui comprend les répétitions A1, A2 et A3, mais des signaux sont aussi obtenus pour le second cadre de lecture qui comprend la séquence de B1. B) Représentation des peptides uniques identifiés dans différentes études de protéomiques à larges échelles. Les bars jaunes représentent l'endroit approximatif des mutations présents dans le peptide comparativement à l'ubiquitine qui permet de les différencier. Ceux-ci sont aussi surlignés en rouge dans la séquence d'acides aminés et les études où ils ont été identifiés sont répertoriées à la fin de chacune des séquences.

### 3.2.4 Modification du protéome par les variants de l'ubiquitine

L'une des propriétés de l'ubiquitine est que celle-ci est capable de s'attacher de manière covalente aux protéines cibles afin de moduler leurs caractéristiques biochimiques et d'influencer leur fonction cellulaire. Cet attachement est possible grâce à la présence de deux glycines en C-terminal de l'ubiquitine. Les protéines A1, A2 et B1 possèdent elles aussi ces deux acides aminés caractéristiques à la fin de leur séquence, ce qui laisse présager que celles-ci pourraient aussi modifier des protéines au sein de la cellule. Pour tester cette hypothèse, nous avons généré des constructions pour exprimer dans des cellules eucaryotes soit l'ubiquitine, soit chacune des protéines d'UBBP4 avec une étiquette HA en N-terminal. Nous avons par la suite transfecté ces constructions dans des cellules HeLa pour 24h et effectuer un immunobuvardage de type western avec un anticorps contre l'étiquette HA.

Pour la piste qui correspond à l'ubiquitine canonique, le signal correspond à une traînée qui débute en haut de la membrane et se termine aux alentours de 20 kDa (Figure 26). Ce type de signal est caractéristique de l'ubiquitine et correspond aux différentes protéines modifiées par celle-ci. Le même phénotype est observé pour les cellules transfectées avec les constructions qui expriment A2 et B1, ce qui démontre que celles-ci ont la capacité de s'attacher à d'autres protéines dans un contexte *in cellulo*. Pour la protéine A3, celle-ci se présente seulement sous forme de monomère qui correspond à la bande de 10kDa visible sur la membrane. Ce résultat était attendu puisque par la faute du décalage du cadre de lecture, les deux glycines ne sont pas présentes et il est donc impossible que cette protéine s'attache de manière covalente à d'autres composantes. De manière surprenante, la protéine A1 est elle aussi retrouvée sous forme de monomère dans la cellule, et ce, malgré la présence dans sa séquence protéique des glycines en C-terminale.



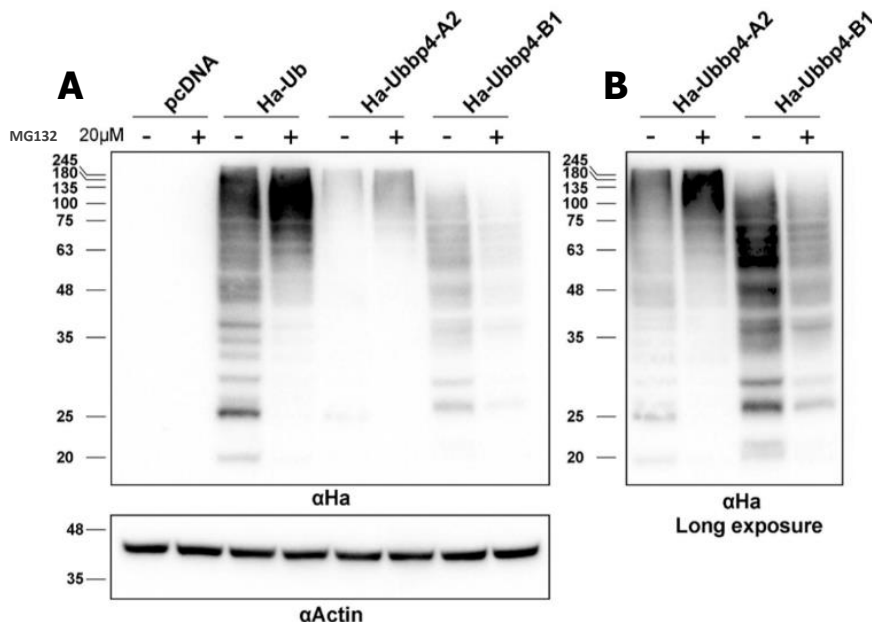
**Figure 26 : Expression des différentes protéines encodées dans le gène UBBP4**

Extraits totaux de cellules HeLa transfecté avec des versions étiquetées HA de l'ubiquitine ou de UBBP4-A1, A2, A3 ou B1. Les extraits ont été séparés sur un gel de SDS, transférés sur membrane et révélés avec un anticorps HA.

### 3.2.5 Réponse d'UBBP4 à l'inhibition du protéasome

Par la suite, nous avons voulu déterminer si les protéines UBBP4-A2 et UBBP4-B1 se comportaient de la même façon que l'ubiquitine canonique en ce qui concerne ses fonctions cellulaires. Pour ce faire, nous avons choisi le rôle le plus connu et le mieux caractérisé pour l'ubiquitine qui est son implication dans l'adressage des protéines à la dégradation par le protéasome. Pour évaluer le rôle de A2 et B1 dans ce mécanisme, nous avons transfecté des cellules HeLa avec les constructions étiquetées HA et nous avons traité les cellules avec un inhibiteur du protéasome, le MG132, pendant 12h. Nous avons par la suite évalué l'effet du traitement par immunobuvardage de type western en utilisant un anticorps dirigé contre l'étiquette HA. Pour les cellules transfectées avec l'ubiquitine, on peut voir que le signal augmente à la suite de l'inhibition du protéasome, ce qui signifie que les protéines modifiées par l'ubiquitine et destinées à la dégradation par ce mécanisme s'accumulent dans la cellule comme attendu (Figure 27). Pour la

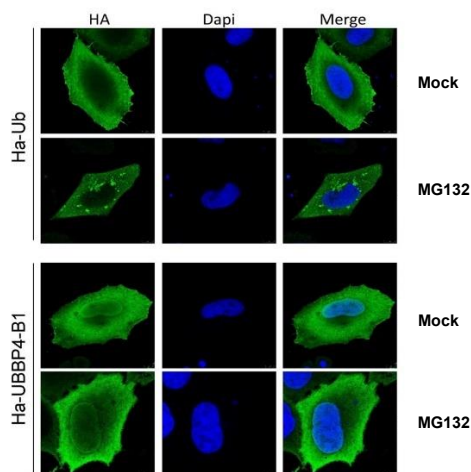
protéine A2, on peut voir que le phénotype est le même que celui de l'ubiquitine canonique avec une augmentation du signal après traitement au MG132 (Figure 27). Cependant, pour les cellules qui expriment la protéine UBBP4-B1, après l'inhibition du protéasome, il n'y a pas d'accumulation de signal, mais au contraire, une disparition de celui-ci tel que le présente la Figure 27. L'expérience a été répétée dans d'autres lignées cellulaires (293 et U2OS) et les résultats obtenus sont semblables à ceux avec les cellules HeLa soit une disparition du signal suite au traitement avec l'inhibiteur du protéasome. Ces résultats laissent présager que UBBP4-B1 modifie les protéines dans la cellule, mais ne récapitule pas la fonction de l'ubiquitine du moins, en ce qui concerne la dégradation protéasomale.



**Figure 27 : Réponse de l'ubiquitine, de UBBP4-A2 et B1 à un inhibiteur du protéasome le MG132**

A) Des cellules HeLa ont été transfectées avec HA-Ub, HA-UBBP4-A2 ou HA-UBBP4-B1, et traitées ou non avec 20μM de MG132, un inhibiteur du protéasome, pendant 12h. Les extraits totaux de ces cellules ont été migrés sur gel de SDS et transférés sur membrane avant d'être hybridés avec un anticorps HA pour visualiser les protéines modifiées ou avec un anticorps actine pour le contrôle de chargement. B) longue exposition pour les transfections avec A2 et B1.

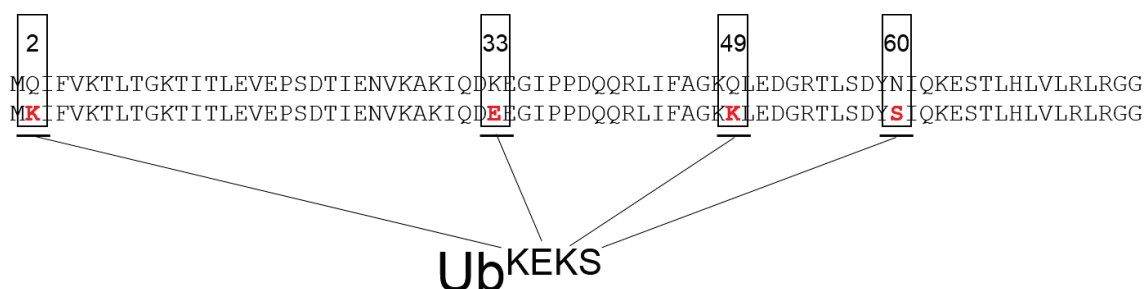
Il est connu que les protéines modifiées par l'ubiquitine s'accumulent dans le cytoplasme sous forme d'agrégats lors de l'inhibition du protéasome avec des agents pharmacologiques (van Eersel et al. 2011). Puisque par immunobuvardage, les protéines modifiées par UBBP4-B1 ne semblent pas être adressées au protéasome, nous avons supposé que ces protéines ne devraient pas non plus former d'agrégats dans la cellule lorsque traitées avec le MG132. Pour y parvenir, nous avons transfectés des cellules HeLa avec l'ubiquitine canonique ou avec UBBP4-B1 et nous avons procédé à une immunofluorescence en utilisant un anticorps contre HA. Tel que rapporté dans la littérature, on voit des agrégats cytoplasmiques dans les cellules exprimant l'ubiquitine canonique et traitées avec le MG132 qui correspond aux protéines dans l'incapacité de se faire dégradées par ce mécanisme. Pour ce qui est des cellules qui expriment UBBP4-B1, on voit bien que la localisation cellulaire des protéines modifiées avant et après inhibition du protéasome n'est pas changée et qu'aucun agrégat n'est présent dans la cellule (Figure 28). Ces résultats apportent une autre preuve que les protéines modifiées par UBBP4-B1 ne sont pas adressées à la dégradation médiée par le protéasome.



**Figure 28 : Immunofluorescence de l'ubiquitine et de UBBP4-B1 à la suite d'un traitement avec un inhibiteur du protéasome**

Des cellules HeLa ont été transfectées avec HA-Ub ou HA-UBBP4-B1 et traitées ou non avec 20µM de MG132 pendant 12h. Les immunofluorescences ont été réalisées en utilisant un anticorps HA pour visualiser la localisation des protéines modifiées par ubiquitine ou par UBBP4-B1.

Par la suite nous avons décidé de renommer la protéine UBBP4-B1 Ub<sup>KEKS</sup> en fonction des 4 acides aminés qui divergent dans sa séquence en acides aminés comparativement à l'ubiquitine. Les mutations présentes au sein de la protéine sont : Q2K, K33E, Q49K, N60S tel que représenté schématiquement à la Figure 29.



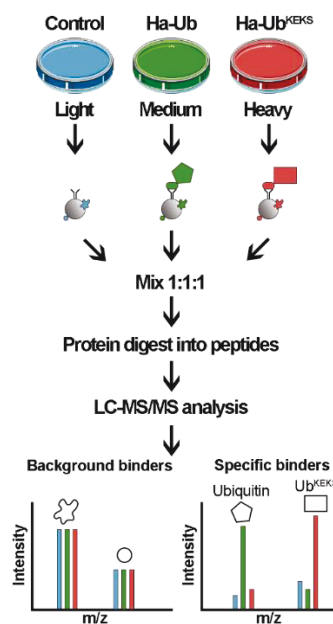
**Figure 29 : Provenance du nom Ub<sup>KEKS</sup>**

UBBP4-B1 a été renommée Ub<sup>KEKS</sup> en fonction des 4 acides aminés différents entre sa séquence et celle de l'ubiquitine.

### 3.2.6 Fonction d'Ub<sup>KEKS</sup>

Puisque Ub<sup>KEKS</sup> ne semble pas posséder les mêmes fonctions que l'ubiquitine, du moins pour la dégradation protéasomale, nous avons voulu identifier le rôle de cette modification post-traductionnelle. Nous avons choisi d'identifier par spectrométrie de masse, les protéines dans la cellule qui sont la cible d'Ub<sup>KEKS</sup> et de voir les fonctions cellulaires qui sont enrichies. Pour y parvenir, nous avons généré des lignées stables qui expriment une répétition de 3 x HA-ubiquitines ou une répétition de 3 x HA-Ub<sup>KEKS</sup>. Nous avons utilisé des cellules HeLa Flip-in et U2OS Flip-in pour générer les lignées stables, ce qui permettait de contrôler à la fois le site d'intégration dans le génome et l'expression des protéines avec la présence du répresseur à la doxycycline. De plus, l'utilisation de ce type de lignée comparativement aux lignées transitoires permet une expression qui est beaucoup moins élevée et qui ressemble plus à un contexte endogène. À la suite de l'obtention de ces lignées, nous avons cultivé les cellules dans du milieu SILAC afin de pouvoir

quantifier nos expériences de spectrométrie de masse. Nous avons donc fait pousser des cellules contrôles (qui n'expriment aucune protéine étiquetée) dans du milieu R0K0, les cellules qui expriment HA-Ub dans le milieu R6K4 et les cellules qui expriment HA- Ub<sup>KEKS</sup> dans le milieu R10K8 tel que schématisé à la Figure 30.



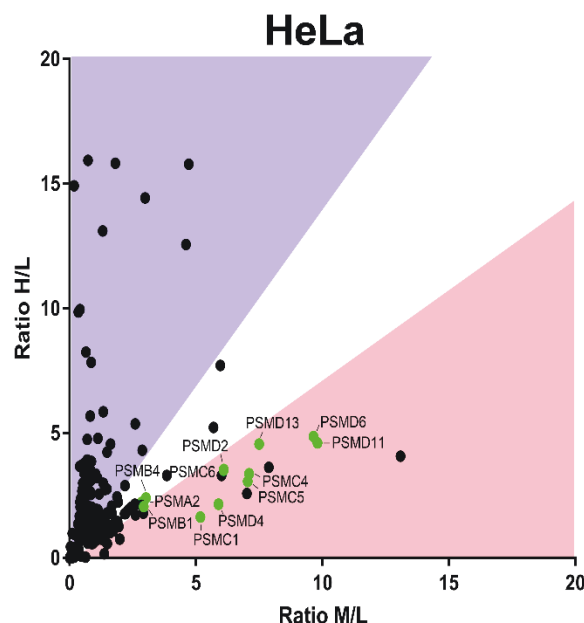
**Figure 30 : Approche expérimentale pour les expériences de spectrométrie de masse**

Les cellules contrôles ont été cultivées dans du milieu léger (R0K0) et les lignées stables ont été cultivées dans du milieu moyen (R6K4) pour l'ubiquitine et dans du milieu lourd (R10K8) pour Ub<sup>KEKS</sup>. Par la suite les protéines provenant des différentes conditions ont été analysées par spectrométrie de masse.

Dans un premier temps nous avons effectué des lyses en condition non dénaturante suivie par une immunoprécipitation contre l'étiquette HA afin d'identifier les protéines qui se retrouvent dans l'environnement de l'ubiquitine et de Ub<sup>KEKS</sup>. Avec ces données, nous avons généré un graphique des ratios M/L (ubiquitine) en fonction des ratios H/L (Ub<sup>KEKS</sup>) pour chacune des protéines identifiées par spectrométrie de masse. Dans la figure 31, les protéines dans la section rouge représentent les protéines plus enrichies avec l'ubiquitine, alors que celles dans la section bleue représentent les protéines qui le sont plus avec Ub<sup>KEKS</sup>. On peut voir que pour l'ubiquitine, plusieurs des protéines enrichies sont des composantes du



protéasome alors que celles-ci le sont beaucoup moins avec Ub<sup>KEKS</sup>. Ces données viennent appuyer notre hypothèse que cette nouvelle modification post-traductionnelle n'adresse pas les protéines pour la dégradation.

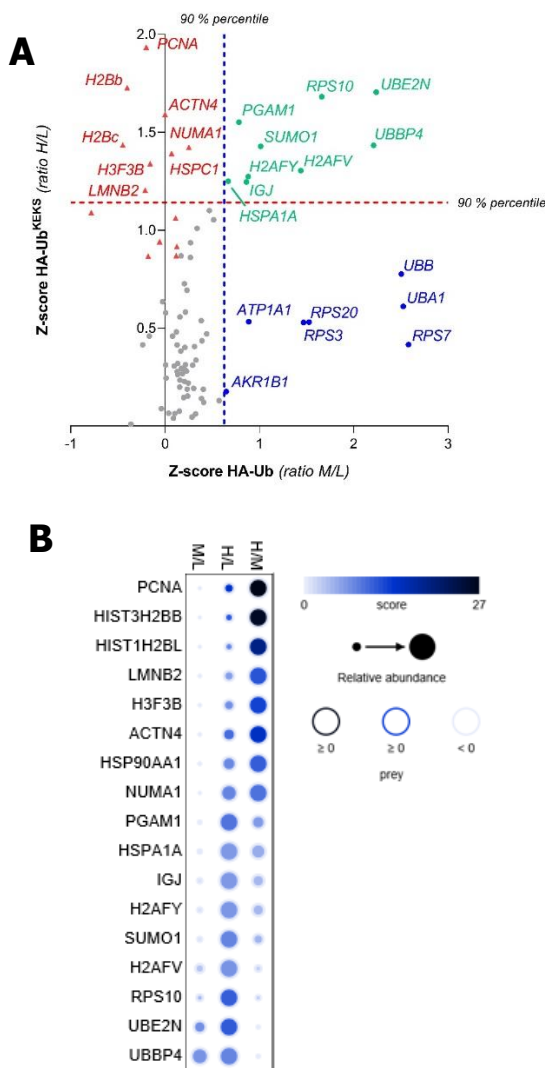


**Figure 31 : Protéines enrichies en conditions natives**

Graphique des protéines qui sont enrichies par immunoprécipitation avec l'ubiquitine (M/L) et avec Ub<sup>KEKS</sup> (H/L). Le triangle rouge contient les protéines qui sont plus enrichies avec l'ubiquitine qu'avec Ub<sup>KEKS</sup> et le triangle bleu délimite les protéines plus présentes avec Ub<sup>KEKS</sup> comparativement à l'ubiquitine. Les protéines du protéasome sont surlignées en vert dans le graphique.

En utilisant les mêmes lignées cellulaires, nous avons refait la même expérience, mais cette fois, en utilisant des conditions de lyse dénaturante afin d'identifier des protéines qui sont modifiées de manière covalente par Ub<sup>KEKS</sup>. Le graphique de la Figure 32 représente les protéines qui sont significativement enrichies avec l'ubiquitine en bleue, les protéines spécifiques à Ub<sup>KEKS</sup> en rouge et les protéines communes en vert. Avec cette approche expérimentale, nous avons trouvé 7 protéines spécifiques à l'ubiquitine et 8 à Ub<sup>KEKS</sup> avec 9 protéines en commun entre les deux. Pour trouver, les meilleurs candidats, nous avons généré un *Dot blot* où chacune des protéines identifiées avec Ub<sup>KEKS</sup> (spécifique ou non) est représentée par un point, plus celui-ci est gros et de couleur foncée, plus la protéine

est enrichie avec Ub<sup>KEKS</sup> (Figure 32). Ainsi en prenant en compte les protéines les plus enrichies avec Ub<sup>KEKS</sup>, on trouve parmi celles-ci la protéine PCNA, des histones et la lamineB2.

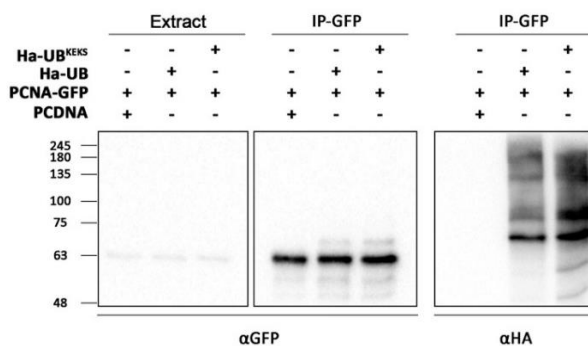


**Figure 32 : Protéines enrichies en condition dénaturante**

A) Graphique des Z-scores pour les ratios M/L (ubiquitine) et (H/L) Ub<sup>KEKS</sup> pour chaque protéine présente dans les trois répliques. Les protéines significativement enrichies (au-dessus du 90% percentile) pour l'ubiquitine sont indiquées en bleue et en rouge pour celles significatives avec Ub<sup>KEKS</sup>. Les protéines en vert sont celles qui sont retrouvées dans les deux conditions. B) Dot plot des protéines qui sont enrichies significativement pour Ub<sup>KEKS</sup>. La grosseur des cercles est proportionnelle au ratio d'enrichissement de la protéine.

### 3.2.7 PCNA-UB<sup>KEKS</sup>

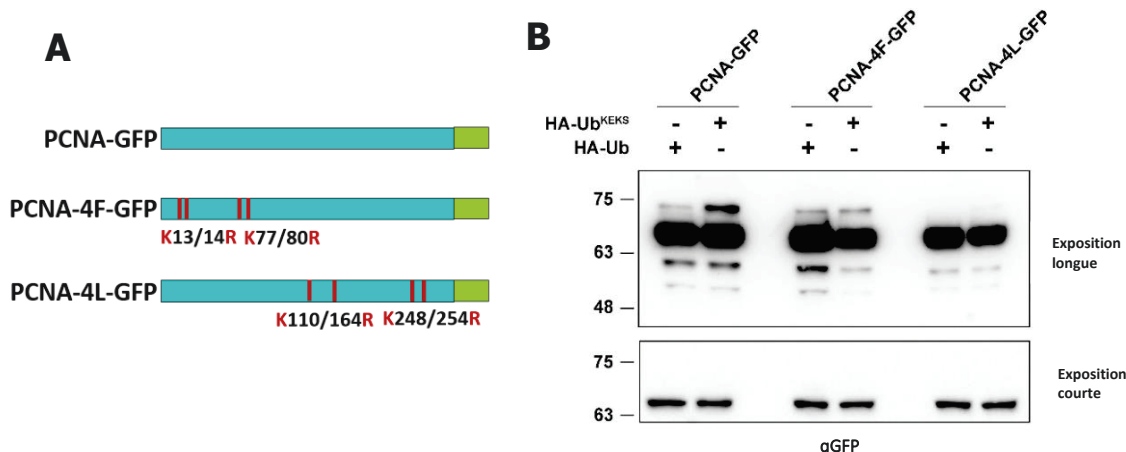
Parmi les protéines qui ressortent significatives avec Ub<sup>KEKS</sup>, celle qui est la plus enrichie en plus d'être spécifique, est la protéine PCNA. PCNA est très bien connue pour son rôle au niveau de la réplication de l'ADN, mais aussi pour son implication dans différents mécanismes de réparations. Le mécanisme de régulation principal qui va déterminer les différentes fonctions de PCNA est via les différentes modifications post-traductionnelles que peut subir la protéine et spécialement l'ubiquitinylation de celle-ci. Puisque PCNA était la protéine la plus enrichie dans nos résultats de spectrométrie de masse et que l'ubiquitinylation de celle-ci était très bien caractérisée, nous avons décidé d'investiguer plus en détail le rôle que Ub<sup>KEKS</sup> pourrait avoir sur cette protéine. Dans un premier temps, nous avons voulu confirmer que PCNA pouvait être modifié par Ub<sup>KEKS</sup>. Pour ce faire, nous avons transfecté des cellules avec PCNA-GFP et HA-Ub ou HA-Ub<sup>KEKS</sup> et nous avons effectué une immunoprécipitation avec un anticorps dirigé contre la GFP qui est fusionnée à la protéine PCNA que nous avons fait migrer sur gel d'acrylamide. Pour visualiser la modification de PCNA, nous avons effectué un immunobuvardage de type western avec un anticorps contre la protéine GFP ou encore pour HA. Comme attendu, la protéine PCNA peut être modifiée par l'ubiquitine telle que le démontre l'apparition d'une bande de plus haut poids moléculaire dans la condition transfectée avec l'ubiquitine avec l'anticorps GFP. Le même profil est visible dans les cellules transfectées avec Ub<sup>KEKS</sup> avec l'anticorps GFP et semble même être plus forte avec cette modification post-traductionnelle comme en témoigne les bandes avec l'anticorps HA. De plus, on peut voir avec l'anticorps HA que la transfection avec les deux modifications mène à l'apparition de plusieurs bandes de plus haut poids moléculaire, suggérant qu'aussi bien l'ubiquitine que Ub<sup>KEKS</sup> peut modifier PCNA sur plusieurs résidus ou encore, former des chaînes sur celle-ci.



**Figure 33 : Modification post-traductionnelle de PCNA par HA-Ub ou de HA-Ub<sup>KEKS</sup>**

Des cellules HeLa ont été transfectées avec PCNA-GFP, HA-Ub ou HA-Ub<sup>KEKS</sup>. Des IP avec des billes GFP ont été effectuées et migrées sur gel de SDS en même temps que les extraits totaux. Les membranes ont soit été révélées en utilisant un anticorps GFP soit avec un anticorps HA.

Puisque Ub<sup>KEKS</sup> est capable de modifier PCNA aussi bien que l'ubiquitine, nous avons voulu investiguer si cette modification se situe sur les mêmes sites qui sont ciblés par l'ubiquitine. Afin de répondre à cette question, nous avons sélectionné 8 sites qui sont connus pour être ubiquitinylés dans des bases de données de spectrométrie de masse. Dans un premier temps, nous avons généré des mutants K/R contenant chacun 4 mutations soit en N-terminal (PCNA-4F) soit en C-terminal (PCNA-4L) (Figure 34). Par la suite, nous avons testé la capacité de ces mutants à être modifiés par Ub<sup>KEKS</sup> en transfectant des cellules avec les mutants étiquetés GFP en présence de HA-Ub ou de HA-Ub<sup>KEKS</sup>. Comme démontré précédemment, la version normale de PCNA peut être modifiée aussi bien par l'ubiquitine que par Ub<sup>KEKS</sup>. Pour ce qui est du mutant 4F, celui-ci semble présenter la même intensité de signal qu'en présence de la forme sauvage lorsque transfecté avec l'ubiquitine, mais l'on constate une diminution de la bande supérieure avec Ub<sup>KEKS</sup> qui correspond à une perte de la forme modifiée de PCNA. Pour le mutant 4L, on perd presque totalement le signal pour la bande supérieure, et ce aussi bien avec l'ubiquitine qu'avec Ub<sup>KEKS</sup>. Ainsi, Ub<sup>KEKS</sup> semble modifier PCNA sur les mêmes résidus que l'ubiquitine et celui-ci semble majoritairement cibler la partie C-terminale de la protéine et plus précisément les lysines 110, 164, 248 et 254.



**Figure 34 : Modification post-traductionnelle des mutants de PCNA par HA-Ub ou par HA-Ub<sup>KEKS</sup>**

A) schéma représentatif des deux mutants générés pour PCNA afin de trouver le site de modification par Ub<sup>KEKS</sup>. B) Des cellules HeLa ont été transfectées avec PCNA-GFP, PCNA-4F-GFP ou PCNA-4L-GFP, et avec HA-Ub ou HA-Ub<sup>KEKS</sup>. Des IP avec des billes GFP ont été effectués et migrés sur gel de SDS. Les membranes ont été révélées en utilisant un anticorps GFP et exposé à différent temps.

### 3.2.8 Délétion du gène d'Ub<sup>KEKS</sup>

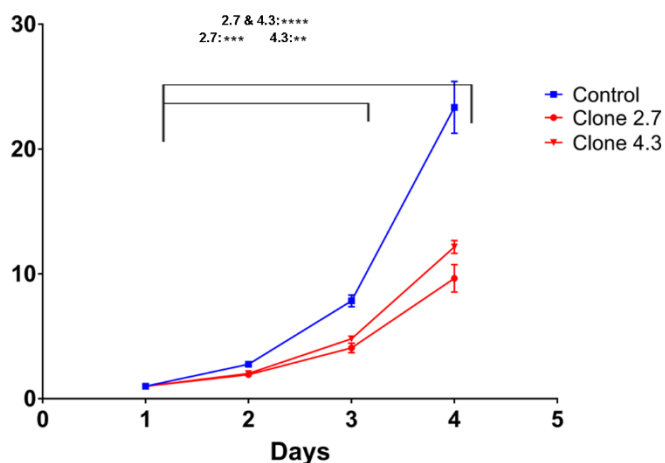
Comme mentionné précédemment l'ubiquitinylation de PCNA est très bien caractérisée pour son rôle au niveau des dommages à l'ADN. Plus précisément, c'est l'ubiquitinylation du site K164 qui est au centre de cette régulation. En fait, la présence d'une mono ou d'une poly-ubiquitinylation sur ce résidu permet l'activation de différents mécanismes de réparation. Nous avons démontré que Ub<sup>KEKS</sup> modifie PCNA sur l'un des 4 sites situés en C-terminale de la protéine, dont le site K164. Afin de déterminer si Ub<sup>KEKS</sup> avait un rôle dans les mécanismes de réparation de l'ADN, nous avons généré des lignées cellulaires dont le gène UBBP4 est délété en utilisant le système CRISPR/Cas9. La figure 35 présente les 2 stratégies utilisées dans les cellules HeLa pour permettre d'enlever le gène complet et ainsi abolir l'expression d'Ub<sup>KEKS</sup>. Puisque nous ne disposons pas d'anticorps pour Ub<sup>KEKS</sup>, nous avons confirmé la délétion du gène par PCR dans un premier temps et par séquençage du génome par la suite.



**Figure 35 : Stratégies utilisées pour les lignées UBBP4<sup>-/-</sup>**

Les lignes mauves représentent la localisation approximative des ARNg utilisés pour enlever le gène de UBBP4. Les deux clones obtenus ont été nommés 2.7 et 4.3 respectivement.

À la suite de la sélection, nous avons choisi deux clones positifs qui présentaient une délétion complète du gène d'UBBP4. Le premier phénotype apparent était que les cellules étaient capables de pousser malgré l'absence de ce gène, ce qui nous laisse penser que celui-ci n'est pas essentiel pour la cellule. Cependant, les cellules sans UB<sup>KEKS</sup> se divisent beaucoup moins rapidement que les cellules HeLa contrôlent. Pour quantifier ce phénomène, nous avons évalué la croissance des cellules en utilisant le CFSE, un marqueur fluorescent qui se divise avec la même vitesse que les cellules que nous pouvons par la suite tracer par FACs et ainsi suivre la croissance des cellules. Tel qu'on peut l'observer dans le graphique de la figure 36, les deux clones de UBBP4<sup>-/-</sup> présentent une courbe de croissance qui est beaucoup plus lente que les cellules contrôles. Cette diminution qui est visible dès le deuxième jour devient significative au troisième jour avec un p-value de 0.001 pour le clone 2.7 et 0.01 pour le clone 4.3. La différence de croissance est flagrante au quatrième jour avec une diminution du taux de croissance de près de la moitié comparativement aux cellules contrôle avec un p-value de 0.0001 pour les deux clones.

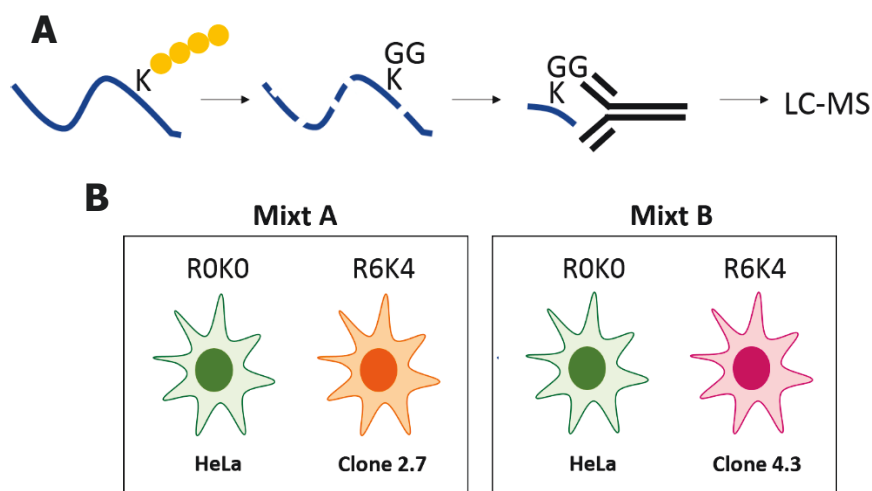


**Figure 36 : courbe de croissance des cellules UBBP4<sup>-/-</sup>**

La courbe en bleue représente la croissance des cellules HeLa contrôles. Les courbes en rouge sont celles des deux clones qui ont une absence d'expression du gène UBBP4. La significativité des résultats est représentée avec les étoiles; \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$  en utilisant un *two-way* ANOVA.

### 3.2.9 Modification du signal di-Glycines dans les cellules Ub<sup>KEKS</sup><sup>-/-</sup>

Afin de déterminer les protéines qui sont ciblées par Ub<sup>KEKS</sup> de manière endogène, nous avons utilisé un anticorps qui reconnaît les deux glycines laissées par l'ubiquitine ou par Ub<sup>KEKS</sup> sur une lysine à la suite d'une digestion avec la trypsine en comparant les marques dans des cellules normales et dans les cellules UBBP4<sup>-/-</sup> (Figure 37). Pour y arriver, nous avons cultivé les cellules contrôles et UBBP4<sup>-/-</sup> dans du milieu SILAC selon les conditions illustrées à la Figure 37. Par la suite, nous avons digéré la même quantité de protéines pour les différentes conditions avec la trypsine et après la purification des peptides nous avons immunoprécipité ceux qui possédaient une marque de di-Glycine et nous avons analysé ces immunoprécipitations par spectrométrie de masse afin de déterminer quels peptides présentaient une diminution de signal dans les cellules <sup>-/-</sup> et donc une perte probable de modification par Ub<sup>KEKS</sup>.

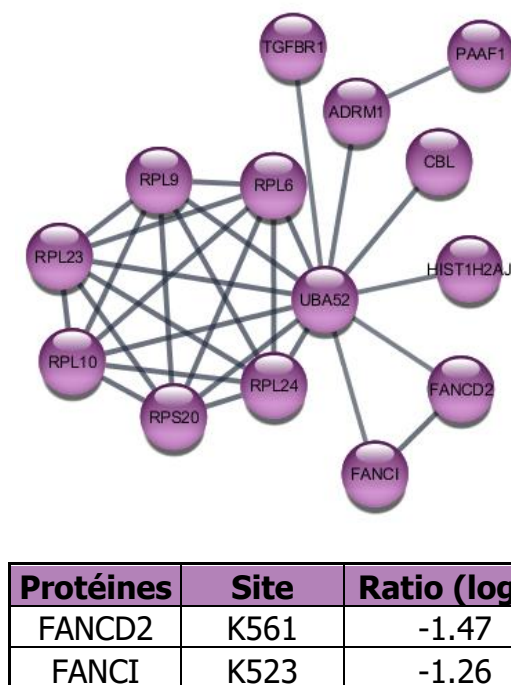


**Figure 37: schéma représentatif de l'approche utilisée pour identifier les cibles endogènes de Ub<sup>KEKS</sup>**

A) Utilisation d'un anticorps spécifique qui reconnaît les motifs de diglycines sur les peptides après digestion avec la trypsine. B) Approche expérimentale pour identifier les sites de modification par Ub<sup>KEKS</sup>.

À la suite de l'analyse des échantillons, nous avons identifié 119 protéines qui présentaient une diminution de la modification diglycine dans le clone 2.7 comparativement aux cellules contrôles. Pour le second clone, 132 protéines présentaient une réduction de signal et seulement 35 de celles-ci étaient communes aux deux lignées cellulaires. Avec ces 35 protéines, nous avons analysé les interactions déjà répertoriées dans la littérature en utilisant la base de données STRING. Avec cette approche, nous avons été capables d'identifier 14 protéines qui interagissaient fortement entre elles (*confidence score cutoff* 0.8). Parmi celles-ci, l'on retrouve plusieurs protéines ribosomales ainsi que deux protéines impliquées dans les mécanismes de réparation de l'ADN et plus particulièrement dans la voie de *Fanconie Anemia*. Un seul site a été identifié chez les deux protéines, soit la lysine 561 pour FANCD2 et la lysine K523 pour FANCI avec une diminution de 1.47 et de 1.26 fois respectivement.





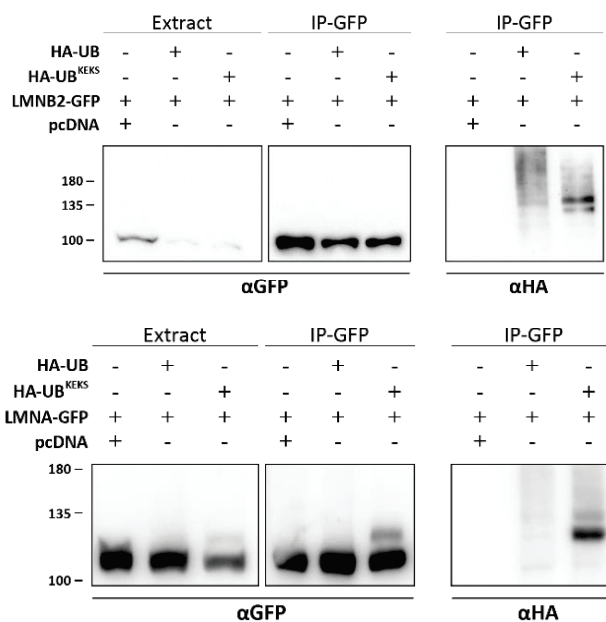
**Figure 38 : Résultats de l'analyse de spectrométrie de masse des protéines avec une diminution de la marque diglycines dans les cellules UBBP4<sup>-/-</sup>**

Schéma des interactions connues pour les protéines communes entre les deux clones avec un score de confiance de 0,8. Le tableau rapporte les lysines identifiées pour FANCI et FANCD2 avec la moyenne des ratios obtenus en log2.

### 3.2.10 Les lamines sont des cibles d'Ub<sup>KEKS</sup>

Parmi les autres cibles d'Ub<sup>KEKS</sup> qui sont significativement enrichies, on retrouve la protéine LMNB2 (Figure 38). De plus, les protéines LMNA et LMNB1 sont aussi retrouvées de manière spécifique dans les IP de Ub<sup>KEKS</sup>, mais ne sont pas significatives dû à leur présence dans seulement 2 des trois répliques. Tout comme pour PCNA, nous avons confirmé la modification de LMNB2 et LMNA par des essais biochimiques. Pour ce faire, nous avons transfecté des cellules avec LMNA-GFP ou LMNB2-GFP en présence soit de HA-Ub ou de HA-Ub<sup>KEKS</sup> et nous avons réalisé une immunoprécipitation des protéines avec des billes GFP. Nous avons par la suite fait migrer ces IP sur gel d'acrylamide et analysé les résultats par immunobuvardage de type western avec des anticorps contre HA ou GFP.

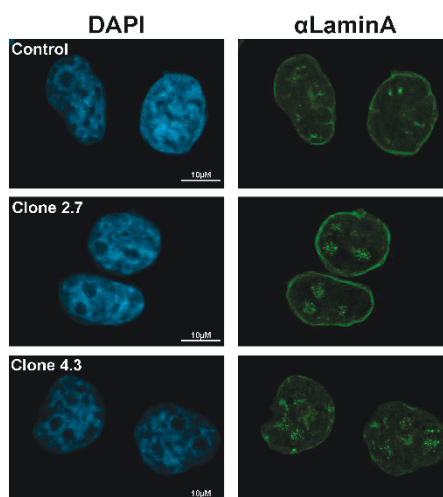
Pour ce qui est de LMNB2 transfecté avec l'ubiquitine, on constate avec l'anticorps HA la présence d'une très faible bande de plus haut poids moléculaire que le poids pour LMNB2-GFP. Cependant, dans les cellules transfectées avec Ub<sup>KEKS</sup>, on peut clairement voir deux bandes de plus haut poids moléculaire, ce qui signifie que LMNB2 peut être modifié par une ou deux molécules d'Ub<sup>KEKS</sup>. Le même résultat a été obtenu lors des essais avec la protéine LMNA-GFP. En effet, lorsque celle-ci est co-transfectée avec Ub<sup>KEKS</sup>, on voit clairement deux bandes qui correspondent à la modification de la protéine LMNA par une ou deux molécules d'Ub<sup>KEKS</sup>. Contrairement à LMNB2 qui présente une très faible bande avec l'ubiquitine, la protéine LMNA semble strictement modifiée par Ub<sup>KEKS</sup> puisqu'on ne constate aucun signal avec l'ubiquitine. Ces deux expériences nous laissent croire que LMNA et LMNB2 sont capables d'être modifiées par la protéine Ub<sup>KEKS</sup> dans un contexte *in cellulo* et viennent donc confirmer ce qui avait été détecté en spectrométrie de masse.



**Figure 39 : Modification post-traductionnelle de LMNB2 et LMNA par HA-Ub ou par HA-Ub<sup>KEKS</sup>**

Des cellules HeLa ont été transfectées avec LMNB2-GFP ou LMNA-GFP et avec HA-Ub ou HA-Ub<sup>KEKS</sup>. Des IP avec des billes GFP ont été effectuées et migrées sur gel de SDS en même temps que les extraits totaux. Les membranes ont soit été révélées en utilisant un anticorps GFP soit avec un anticorps HA.

Puisque Ub<sup>KEKS</sup> n'est pas impliquée dans la dégradation des protéines par le protéasome, nous avons voulu explorer la fonction de cette modification post-traductionnelle sur les protéines LMNB2 et LMNA. Pour ce faire, nous avons effectué des immunofluorescences des protéines endogènes dans les cellules UBBP4<sup>-/-</sup> en utilisant des anticorps spécifiques pour chacune des protéines. Il s'est avéré que LMNB2 présentait une localisation similaire entre les cellules contrôles et les cellules sans l'expression d'Ub<sup>KEKS</sup> (résultat non montré). Cependant, la protéine LMNA, bien que toujours présente à l'enveloppe nucléaire comme dans les contrôles, présente une sous-population qui se localise dans les nucléoles, et ce, dans les deux clones étudiés. Ces résultats nous laissent présager une quelconque fonction d'Ub<sup>KEKS</sup> dans la localisation cellulaire de la protéine LMNA.

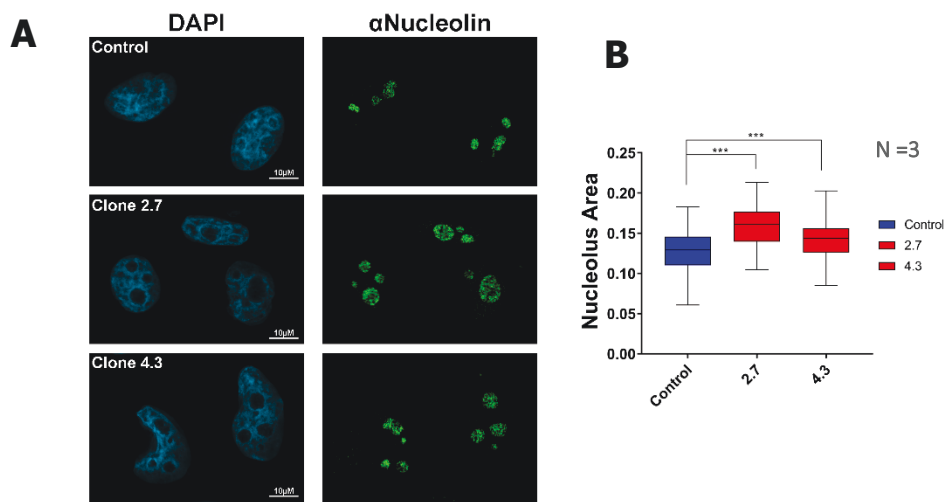


**Figure 40 : Localisation de LMNA dans les cellules UBBP4<sup>-/-</sup>**

La localisation cellulaire de LMNA endogène a été évaluée en effectuant des immunofluorescences dans les cellules UBBP4<sup>-/-</sup> et dans les cellules contrôles. Les noyaux ont été colorés à l'aide du DAPI et LMNA a été visualisés en utilisant un anticorps dirigé contre la protéine. Les images ont été prises avec un microscope confocal et une seule coupe est présentée ici.

Un autre phénotype évident que nous avons remarqué en regardant la localisation de LMNA est que les cellules UBBP4<sup>-/-</sup> présentent des nucléoles de plus grande dimension comparativement aux cellules HeLa contrôles. Tel que le démontrent les immunofluorescences de la Figure 40, on peut clairement voir que les nucléoles sont

beaucoup plus gros (représenté par les trous noirs dans le DAPI) dans les cellules sans Ub<sup>KEKS</sup> en comparaison des cellules HeLa de type sauvage. Pour quantifier ce phénotype, nous avons utilisé un marqueur des nucléoles, la nucléoline, et nous avons calculé l'aire occupée par ceux-ci au sein du noyau en utilisant le DAPI comme référence de l'aire totale du noyau. Le graphique représenté à la figure 41 nous indique que les nucléoles dans les cellules UBBP4<sup>-/-</sup> occupent un espace significativement différent comparativement à des cellules contrôles.



**Figure 41 : Quantification de l'aire nucléoles dans les cellules UBBP4<sup>-/-</sup>**

Les nucléoles des cellules HeLa contrôles et les cellules UBBP4<sup>-/-</sup> ont été marqués par un anticorps contre la protéine nucléoline en vert et les noyaux en bleu ont été marqués par le DAPI et visualisés par un microscope confocal. B) En utilisant un logiciel (CellProfiler), nous avons quantifié l'aire des nucléoles par rapport à celle du noyau en utilisant le DAPI comme marqueur de ceux-ci. En bleu, on peut voir l'aire des nucléoles dans les cellules contrôles, et en rouge, l'aire de ceux-ci dans les deux clones UBBP4<sup>-/-</sup> avec la moyenne indiquée pour chacune des conditions. Un Two-way ANOVA a été effectué comme test statistique et les trois étoiles (\*\*\*) signifie un  $p < 0.001$ .

## **4 Discussion**

### **4.1 Chapitre 1 : Étude des cibles du complexe DDB1-CulA via l'intermédiaire des protéines DCAFs**

Les complexes d'ubiquitines ligases sont le sujet d'innombrables publications scientifiques depuis de nombreuses années. Au cœur de presque tous les processus cellulaires, ce sont ces complexes qui ont la responsabilité d'identifier les protéines cibles qui doivent être ubiquitinylées. Puisque la cellule ne peut pas posséder le même nombre de complexes d'ubiquitines ligases que le nombre de protéines à être ubiquitinylées, cette dernière a mis en place des stratégies pour qu'un seul complexe puisse reconnaître une grande variété de protéines. Cette méthode est utilisée par les complexes de Cullins (CRLs) qui vont tous posséder une structure de base identique comportant des protéines d'échafaudages (figure 9) qui vont s'associer avec différentes protéines responsables de la reconnaissance des substrats. Ainsi, un complexe peut cibler une multitude de protéines et intervenir dans plusieurs processus cellulaires. C'est par l'utilisation de ces nombreuses protéines de reconnaissance que le complexe DDB1-Cul4 peut intervenir dans des processus cellulaires tels que la transcription, le contrôle du cycle cellulaire, la réplication et la réparation de l'ADN (Sang, Yan, et Ren 2015). Ces protéines de reconnaissance portent le nom de DCAF et forment une famille qui compte près de 60 membres.

Cette famille possède la particularité d'avoir un domaine caractéristique dérivé des domaines WD40 traditionnels. Ce domaine est important pour l'association avec la protéine DDB1. Les domaines WD40 classiques se composent d'un minimum de quatre répétitions WD40 (*WD40 repeats*) qui sont constituées d'environ 40 acides aminés et se terminent généralement par un tryptophane (W) et une aspartate (D), d'où leur nom WD40. Les DCAFs présentent un domaine WD40 légèrement modifié

avec une séquence WDXR40 qui a été découverte par l'analyse des DCAFs qui étaient déjà connues. Par la suite, l'utilisation de ce domaine a permis d'identifier par bio-informatique plusieurs autres protéines avec ce type de séquences (J. Jin et al. 2006).

Contrairement aux autres complexes CRL présents dans la cellule, l'interaction entre les DCAFs et la protéine DDB1 présente une très large flexibilité dans le mode de liaison (Petroski et Deshaies 2005b). En fait, des études ont démontré que pour certaines des DCAFS, des résidus en dehors des domaines WDXR40 sont nécessaires pour leur interaction avec la protéine DDB1 (J. Jin et al. 2006). Ceci indique que la seule présence du domaine n'est pas nécessairement associée à une interaction avec DDB1. De plus, il a été démontré que des protéines virales (tel que SV5 V) sont capables de lier la protéine DDB1 sans toutefois présenter un domaine WDXR40 dans leur séquence (Precious et al. 2005). Ainsi, la liaison de la protéine DDB1 avec les différentes DCAFs présente un large éventail de possibilités et on ne peut donc pas se fier strictement à la présence du domaine WDXR40.

Les protéine DCAFs ont initialement été bâties sur des études de spectrométrie de masse qui utilisaient seulement les membres du complexe CRL4 (DDB1 ou CUL4) pour définir leur appartenance à cette famille protéique (J. Jin et al. 2006). De plus, l'identification de la présence d'un domaine WDXR40 n'est pas une preuve irrévocable de l'association d'une protéine avec la protéine DDB1. D'autres protéines ont aussi été découvertes pour interagir avec ce complexe d'ubiquitine-ligase après que la liste initiale ait été publiée ce qui fait qu'elles ne se retrouvent pas dans cette dernière (Ohno et al. 2010; Zeng et al. 2016). Finalement, pour certaines de ces protéines, aucune étude n'a été faite quant à la nature de leur fonction au sein de la cellule. Puisque le complexe CRL4 est essentiel dans de nombreux processus cellulaires, il est important de bien connaître toutes les protéines qui sont responsables de l'identification de ses substrats afin de mieux comprendre la régulation apportée par celui-ci. Pour ces différentes raisons, nous avons décidé de

nous servir des protéines DCAFs, au lieu de seulement utiliser la protéine DDB1, pour mieux caractériser les complexes formés par DDB1-Cul4.

#### **4.1.1 Profil des protéines DCAFs**

Le complexe CRL4 est impliqué dans la dégradation de nombreuses protéines clés qui régulent différents processus cellulaires de manière coordonnée. Par exemple, le complexe DDB1-Cul4-DCAF2 (aussi connu sous le nom de Cdt2 ou DTL) est essentiel pour la progression de la phase S puisqu'il permet la dégradation de protéines importantes telle que Cdt1, ce qui empêche la re-réplication de l'ADN (Sheng et al. 2019). Il n'est donc pas étonnant de constater que dans les résultats des analyses des protéines essentielles, la protéine DCAF2 ressort comme très importante dans la croissance cellulaire (Figure 11). Il était aussi attendu que la protéine DDB1 soit parmi celles indispensables à la survie cellulaire. Puisque c'est elle qui est au cœur de la formation de tous les complexes, l'abolition de son expression entraîne la dérégulation de nombreux mécanismes et donc l'arrêt de la croissance. Pour ce qui est des protéines Cul4a et Cul4b, les cellules supportent très bien l'absence de ces deux protéines. Il avait déjà été démontré qu'il y avait une certaine redondance entre les fonctions de ces deux protéines puisque dans des études ultérieures, leur délétion ne semblait pas affecter de manière significative les cellules (Hannah et Zhou 2015).

Dans cette étude, nous avons identifié que 19 protéines DCAFs altéraient de manière significative la croissance cellulaire lorsque l'expression de celles-ci était diminuée. Parmi celles-ci, on retrouve des protéines bien connues tel que DCAF2 et DCAF1, mais aussi des protéines qui n'ont pas encore été largement étudiées. Par exemple, la protéine WDR57 (nommé aussi SNRNP40) a été répertoriée pour être associée au complexe U5 du spliceosome (Jurica et al. 2002). Le spliceosome est constitué de plusieurs sous-complexes formés de protéines et d'ARN (nommé U1 à U6) qui vont séquentiellement agir sur le pré-ARNm pour enlever les introns de sa séquence afin de produire une molécule d'ARNm mature. Chacun des sous-complexes possèdent

son environnement protéique spécifique qui est nécessaire au bon déroulement du processus d'épissage alternatif. La protéine WDR57 fait partie du sous-complexe U5 qui est impliqué au niveau de la fusion entre deux exons (Will et Lührmann 2011). Mis à part son rôle dans la formation de ce complexe, la fonction précise de la protéine WDR57 n'a pas été étudiée. Il est intéressant de mentionner que le spliceosome est un complexe hautement dynamique composé de dizaines de protéines qui vont s'interchanger durant les différentes étapes du processus. Le fait que la protéine WDR57 soit essentielle à la survie cellulaire met en perspective le fait que celle-ci doit être importante dans la régulation d'une des étapes de ce processus cellulaire. Il a été démontré que l'ubiquitine peut jouer un rôle important dans la formation du spliceosome dans la cellule (Bellare et al. 2008). Il se pourrait que WDR57 puissent cibler certaines protéines du complexe U5, afin de permettre leur ubiquitinylation et donc permettre l'assemblage d'un complexe fonctionnel (Bellare et al. 2006). Par exemple, il a été démontré que la protéine Prp8, une des protéines qui forme le complexe U5, est ubiquitinylée *in vitro* et interagit avec WDR57 dans nos IPs. Il se pourrait que la protéine Prp8 soit modifiée par l'ubiquitine via le complexe DDB1-Cul4 par l'intermédiaire de WDR57.

Il n'est pas impossible que certaines des DCAFs aient échappé au décompte de protéines essentielles par la faute des critères établis ou encore par le choix des lignées cellulaires dans laquelle l'étude a été faite. Bien que le but soit d'identifier les protéines essentielles dans tous les types cellulaires, il se pourrait que certaines soient indispensables à la croissance cellulaire dans un contexte particulier. De plus, cette étude a été faite à large échelle, ce qui peut remettre en cause l'identification de protéines essentielles. En fait, dans des expériences visant un autre objectif, nous avons tenté de générer des shARN contre la protéine DCAF13 et nous avons noté une diminution importante de la croissance des cellules lorsque l'expression de la protéine était diminuée. Le phénotype étant tellement important que nous n'avons pu continuer les expérimentations prévues initialement avec ces cellules puisque celles-ci ralentissaient de manière importante leur croissance (résultats non



montrés). Pourtant cette dernière n'apparaît pas dans la liste des protéines DCAFs nécessaires à la viabilité cellulaire. Il est donc probable que toutes les DCAFs qui possèdent cette particularité n'aient pas été identifiées dans cette étude. Néanmoins, les informations que nous avons pu trouver grâce à cette étude nous donnent des renseignements utiles pour la construction des profils de chacune des protéines DCAFs.

#### **4.1.2 Partenaires d'interaction**

##### **4.1.2.1 Immunoprécipitations de la protéine DDB1**

Les études de spectrométrie de masse qui ont répertorié les protéines DCAFs ont toutes utilisé la protéine DDB1 ou les protéines Cul4a/b comme appât (Angers et al. 2006; He et al. 2006) . Dans la majorité des cas, les interactions ont été confirmées dans un dans un système avec transfection transitoire où les 3 composantes protéiques du complexe (DCAF, DDB1 et Cul4) étaient surexprimées. De plus, toutes ces études n'ont pas utilisé de méthodes quantitatives ou statistiques (tel que le SILAC ou des méthodes d'analyse telles que le logiciel SAINT) pour déterminer l'interaction entre les composantes du complexe CRL4 et les DCAFs. Ceci peut remettre en question leur appartenance à cette famille. Quelques protéines DCAFs sont déjà bien connues et caractérisées pour leur rôle au sein du complexe CRL4, comme par exemple DCAF1 et DCAF3. Cependant, comme mentionné, certaines n'ont été vues que dans une des études et n'ont pas été reconfirmées par la suite. Afin de construire un profil du complexe DDB1 le plus fidèle possible, nous avons utilisé à la fois la protéine DDB1 et les protéines DCAFs.

Pour les IPs avec la protéine DDB1, nous avons combiné une approche GFP avec une approche BirA\* pour augmenter nos chances d'identifier des protéines DCAFs. L'utilisation de SILAC nous a permis de quantifier ces interactions afin de déterminer l'enrichissement de chacune des DCAFs. Pour ces deux méthodes, nous avons

identifié 20 protéines DCAF avec l'approche GFP et 15 avec la protéine de fusion BirA\*. Ces résultats ressemblent beaucoup à ceux obtenus dans les autres articles de recherche qui se sont intéressés à la question, puisque comme mentionnés, le même type d'approche fut utilisée pour l'identification des DCAFs (Angers et al. 2006; Shiyonov, Nag, et Raychaudhuri 1999). Cependant, contrairement à ce que nous nous attendions, très peu de protéines WDR ont été identifiées par cette approche. En fait, l'étude qui a présenté les interactions entre plusieurs des protéines WDR et le complexe DDB1 a utilisé l'homologie du domaine pour identifier ces protéines et a par la suite surexprimé les composantes afin de vérifier leur interaction. Ceci nous laisse croire que la surexpression a peut-être entraîné des interactions qui, de manière endogène, ne sont pas présentes dans la cellule. Il se pourrait cependant que les interactions nécessitent un contexte cellulaire particulier pour interagir avec le complexe CRL4, ce qui limite notre capacité d'identification. Toutefois, le fait que ces protéines n'ont pas été aperçues dans les autres études de spectrométrie de masse nous amène à penser que ces protéines n'appartiennent pas à la famille des DCAFs. Un autre aspect surprenant est la quantité de protéines identifiée avec l'approche BirA\*. En fait, nous nous serions attendus à plus de protéines DCAFs avec l'approche BirA\* qu'avec celle du GFP puisque celle-ci permet d'identifier des protéines sur une période de temps de 24h comparativement à l'approche GFP où les interactions protéiques sont figées au moment de l'IPs.

#### **4.1.2.2 Confirmation de l'interaction des protéines DCAFs avec le complexe DDB1-Cul4**

Parmi la famille des DCAFs qui compte 59 membres, nous avons obtenu l'ADNc et généré les lignées cellulaires pour 57 d'entre elles (nous n'avons pas été en mesure d'amplifier les protéines DCAF1 et DCAF19) (J. Lee et Zhou 2007b). Sur ces 57 protéines, les IPs ont été réalisées sur 50 d'entre elles (des problèmes rencontrés lors de la réalisation des expériences de spectrométrie de masse font que les IPs sur les protéines TLE1, TLE2, WDR12, WDR59, DCAF17, DCAF5 et DCAF2

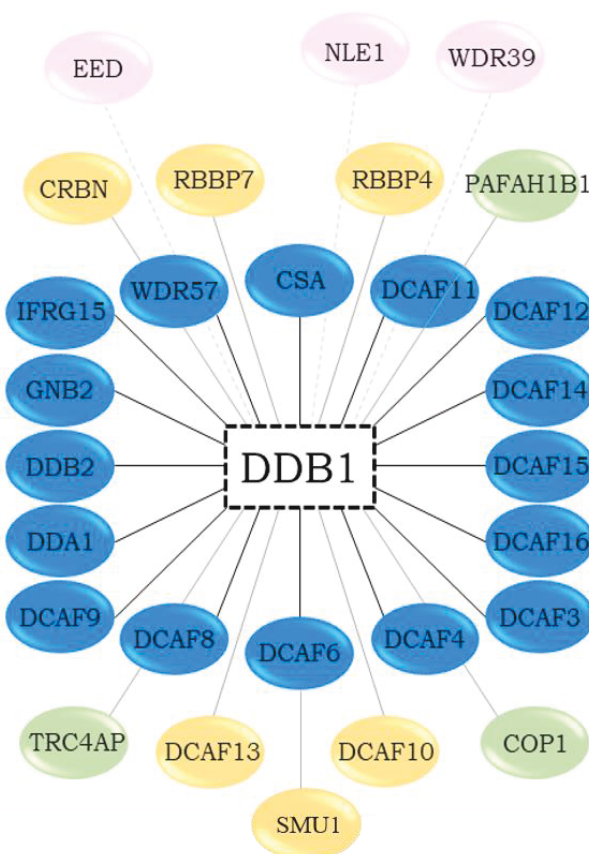
devront être refaites). Afin d'identifier toutes les protéines qui se retrouvent dans l'environnement de la DCAF en question, nous avons utilisé l'approche BirA\* telle que décrite précédemment. Nous avons utilisé deux types de contrôles dans nos IPs : le premier étant des cellules vides et le second des cellules exprimant la protéine BirA\* seule. Ces deux contrôles nous permettent d'identifier les protéines qui s'associent aux billes de streptavidine de manière non spécifique et les protéines qui sont modifiées de façon non spécifique par BirA\*. Au cours de l'étude, nous nous sommes aperçus que l'utilisation du second type de contrôle (BirA\* seule en présence de biotine) amène l'identification d'un très grand nombre de protéines puisque BirA\* est "libre" dans la cellule. Ceci entraîne la biotinylation de protéines aléatoires dans la cellule et possiblement de celles qui sont le plus abondantes. Un exemple de ce phénomène dans nos résultats peut être illustré par la protéine Cul4b. Celle-ci a été identifiée dans le contrôle BirA\* seul avec 27 et 41 comptes MS/MS. Ce résultat a pour conséquence que plusieurs protéines DCAFs, qui ont aussi plusieurs comptes avec Cul4b, ont obtenu un faible score SAINT (allant même jusqu'à zéro) pour des protéines DCAFs qui ont été répertoriées dans la littérature pour avoir une interaction avec le complexe. C'est le cas entre autres de la protéine DCAF12 qui a obtenu 1|0 compte pour la protéine Cul4a et 26|34 pour Cul4b, ce qui a tout de même donné un score SAINT de 0 pour les deux protéines. Toutefois, l'interaction entre la protéine DCAF12 et Cul4 a été répertoriée dans la littérature (Ravichandran et al. 2019). Ainsi, pour certaines protéines, le contrôle BirA\* est trop sévère pour permettre aux interacteurs d'obtenir un score suffisant pour être significatif. Ce n'est cependant pas le cas pour les protéines DDB1 et Cul4a qui ont obtenu dans les contrôles un nombre de comptes raisonnable qui a permis de bien scorer les protéines. Un contrôle idéal serait d'utiliser une protéine de fusion telle que GFP-BirA\* avec ou sans signal de localisation cellulaire et d'utiliser celui qui ne possède pas la même localisation cellulaire que notre protéine d'intérêt.

Si l'on résume les résultats obtenus avec les IPs DCAFs en ce qui concerne les protéines DDB1-Cul4, 16 DCAFs ont présenté une interaction significative avec

au moins un des membres du complexe et seulement 3 qui ont des scores non significatifs. De plus, si l'on considère que les scores de Cul4b ont peut-être été faussés par l'utilisation du contrôle BirA\* seul, 21 protéines DCAFs ont présenté un minimum de 20 comptes MS/MS avec celle-ci. Ceci résulte, par exemple, à l'exclusion de DCAF12 dans les protéines de cette famille dont nous sommes certains qu'elles interagissent avec le complexe DDB1-Cul4.

Les résultats obtenus avec les IPs DDB1 et celles de chacune des DCAFs ont été regroupés et sont présentés à la figure 42. Les protéines DCAFs ont été disposées en fonction de leur niveau de confiance par rapport à leur interaction avec le complexe. Les protéines en bleu sont celles qui ont été retrouvées dans les IPs avec la protéine DDB1, et qui dans leurs IP respectives, elles ont obtenu une interaction significative avec une ou plusieurs des composantes du complexe CRL4 (score SAINT  $\geq 0,7$ ). Les protéines au second niveau ont démontré des interactions avec la protéine DDB1 (Jaune) ou avec les composantes du complexe dans leur propre purification par affinité (Vert) (toujours avec un score de  $\geq 0,7$ ). Finalement en rose, ce sont des protéines qui ont obtenu un score SAINT inférieur à 0,7 avec les protéines DDB1-Cul4, mais qui ne sont tout de même pas négatives. Ainsi, cette liste comporte les protéines qui font réellement partie de la famille des DCAFs en condition physiologique. Notre liste ressemble beaucoup à celle qui avait été trouvée par les études publiées précédemment avec quelques ajouts en plus (tel que la protéine WDR57) (J. Lee et Zhou 2007b). Le fait que de nombreuses protéines WDR n'aient pas été retrouvées ni avec DDB1 ni dans les IPs des DCAFs diminue grandement la liste des protéines qui peuvent s'associer au complexe. Ceci fait que la protéine DDB1 posséderait un nombre beaucoup plus limité de protéines adaptatrices que ce qui avait été suggéré initialement. Il est possible que certaines des protéines de la liste initiale qui n'ont pas été concluantes dans cette étude puissent tout de même former un complexe avec les protéines DDB1-Cullin4. En effet, si la protéine interagit avec le complexe de façon minoritaire ou encore dans certaines conditions physiologiques précises (ex : type de dommages particuliers ou

voie métabolique spécifique), il est probable que nous n'avons pas été en mesure de démontrer leur interaction. De plus, parmi les protéines manquantes dans nos expériences, certaines pourraient faire partie de la famille des DCAFs, par exemple la protéine DCAF1 qui a été largement caractérisée pour son rôle dans la réponse virale en association avec DDB1 (Schabla, Mondal, et Swanson 2018).



**Figure 42 : Réseau d'interaction des protéines avec DDB1 selon leur degré de confiance**

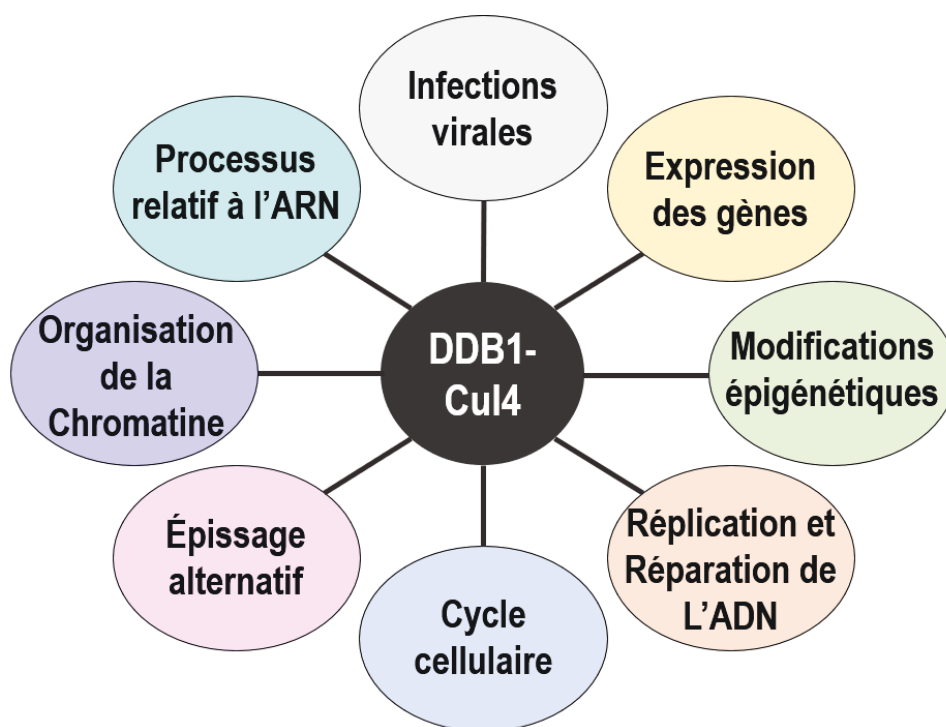
Schéma des interactions pour la protéine DDB1 et les protéines DCAFs. En bleu ce sont les interactions qui ont été rapportées à la fois dans les IPs avec DDB1 et avec celles des DCAFs de manière individuelle. En jaune, les protéines DCAFs qui ont été vues seulement dans les IPs DDB1, mais sans confirmer la réciproque. En vert, ce sont celles qui ont identifié l'une des composantes du complexe CRL4 dans leur IPs, mais qui n'ont pas été vues avec la protéine DDB1. Finalement en rose, ce sont les protéines qui ont identifié des protéines du complexe CRL4, mais qui n'ont pas obtenu un score SAINT assez élevé pour confirmer hors de tout doute leur interaction.

#### **4.1.2.2 Identification des fonctions principales des protéines DCAFs.**

L'analyse des partenaires d'interactions des protéines DCAFs nous permet certes de nous renseigner sur la formation des complexes avec DDB1, mais nous permet aussi de déterminer, via les protéines identifiées dans les IPs, les fonctions cellulaires de chacune d'entre elles. On peut donc ainsi étendre le rôle des protéines DCAFs au complexe CRL4 pour déterminer ses diverses fonctions cellulaires. La figure 16 représente les processus cellulaires qui ont été significativement enrichis par les protéines DCAFs. Bien qu'on dénombre plus d'une trentaine de fonctions, on peut les regrouper en 8 fonctions principales qui sont présentées à la figure 43. Plusieurs des fonctions retrouvées ont déjà été associées au complexe DDB1-Cul4 telles que le cycle cellulaire, la réplication et la réparation de l'ADN. Ce qui nous indique que nos résultats concordent avec la littérature.

Par exemple, la protéine COP1 (aussi connu sous le nom de RFWD2) a été décrite dans la littérature pour être impliquée au niveau de la dégradation de protéine telle que p27 et p53 (Ka et al. 2018). En fait, il a été démontré que suite aux dommages à l'ADN, la protéine COP1 est phosphorylée par ATM pour promouvoir sa dégradation, ce qui stabilise la protéine p53 (Dornan et al. 2006). Parmi les protéines que nous avons identifiées avec COP1, on retrouve un enrichissement significatif de protéines impliquées dans le cycle cellulaire, ce qui concorde avec la littérature. Cependant, le fonctionnement de la protéine COP1 n'est pas encore très bien caractérisé. Parmi les protéines que nous avons identifiées dans nos IPs, nous retrouvons plusieurs protéines qui ont un lien direct avec la réplication de l'ADN. Ceci nous laisse croire que la protéine COP1 pourrait avoir une fonction lors de la réplication de l'ADN et que lors de dommages, celle-ci serait dégradée pour permettre la stabilisation de la protéine p53. Un aspect intéressant de la protéine COP1 est qu'elle possède un domaine RING capable de recruter des E2 pour promouvoir leur ubiquitinylation (Dornan et al. 2004). Cependant, l'activité de la protéine COP1 seule ou encore en complexe avec la protéine DDB1 ne semble pas présenter les mêmes fonctions cellulaires (Ly et al. 2015). Afin de déterminer les fonctions de la protéine DDB1 avec le complexe CRL4, il faudrait muter son domaine RING pour identifier les protéines qui sont ciblées de manière différentielle.

Parmi les fonctions que nous avons retrouvées et qui jusqu'à présent n'ont pas été étudiées dans le contexte de la formation du complexe DDB1-Cul4 est celle associée à l'épissage alternatif. Une seule étude chez la plante a démontré que des protéines associées au spliceosome pouvait interagir avec le complexe DDB1-Cul4, mais cette étude n'a fait état que de l'interaction sans toutefois investiguer les fonctions de celle-ci (J.-H. Lee et al. 2008). Parmi les DCAFs, les protéines WDR57, DCAF4 et DDA1 sont celles qui présentent des interactions avec les protéines associées à cette fonction. La protéine WDR57, présentée précédemment dans le travail, est connue pour être une composante protéique du spliceosome. La protéine DCAF4 cependant n'a jamais eu de fonction rapportée dans la littérature outre quelques études de corrélation entre son expression et la fréquence de certains cancers (H. Liu et al. 2017). La mise en évidence que DCAF4 peut s'associer avec des protéines impliquées dans l'épissage alternatif vient proposer pour la première fois une fonction de celle-ci.



**Figure 43: Fonctions du complexe DDB1-Cul4**

Principales fonctions identifiées pour le complexe DDB1-Cul4 via les protéines DCAFs.

### **4.1.3 Identification des cibles des protéines DCAFs.**

L'une des conséquences de l'interaction entre un complexe d'ubiquitine ligase et sa cible est que cette dernière est adressée, dans la majorité des cas, à la dégradation via le protéasome. Ceci a comme conséquence l'impossibilité d'identifier la protéine que ce soit par des méthodes biochimiques ou en spectrométrie de masse. Afin d'identifier les substrats des protéines DCAFs identifiées, nous avons opté pour une approche de Pulse-SILAC afin de mesurer la différence de ratio entre des cellules normales et des cellules dont les protéines DCAFs sont surexprimées. Chacune des protéines DCAFs présente donc des cibles qui sont différentes et nous sommes capables de les identifier en utilisant la différence d'abondance des protéines dans des cellules normales ou dans celles qui surexpriment une DCAF. L'un des problèmes avec cette approche que nous avons identifiés par la suite est que celle-ci ne nous permet malheureusement pas de discriminer les protéines qui diminuent en abondance de celles qui sont surexprimées. En effet, dans la plupart de nos expériences la protéine avec le plus haut ratio H/M est la dite DCAF en question. De plus, le complexe DDB1 est largement impliqué dans les mécanismes de régulation de la transcription autant négativement que positivement. Ainsi, l'augmentation ou la diminution d'un transcrit peut résulter par un changement de l'abondance de la protéine qui se reflète dans nos expériences de spectrométrie de masse. Il reste cependant difficile de discuter des résultats obtenus pour chacune des protéines à l'étude, que ce soit pour l'identification de leur cible aussi bien que pour leurs partenaires d'interaction. Pour cette raison, nous avons sélectionné une protéine en particulier qui a fait l'objet d'une étude plus détaillée.

### **4.1.4 DDB2 et TCERG1**

La protéine DDB2 est un bon exemple de l'ensemble des résultats que nous avons obtenus. Dans ce cas-ci, les données d'ARNm et les résultats de croissance cellulaire ne nous renseignent pas immédiatement sur l'importance de DDB2. Néanmoins, on peut constater que les cellules expriment son ARNm, mais qu'elles semblent capables de survivre en son absence. Ceci signifie que son rôle ne touche



pas ou peu aux fonctions qui maintiennent la croissance cellulaire. Cette interprétation concorde avec la littérature puisque la fonction principale de la protéine DDB2 est d'assurer le maintien de l'intégrité génomique via les mécanismes de réparation de l'ADN. Bien que ces processus soient nécessaires dans un organisme complexe, les cellules en culture sont capables de survivre en son absence, ce qui est le cas lors de la délétion du gène de la protéine DDB2.

Les IPs de DDB2 ont permis d'enrichir significativement 24 protéines si l'on considère seulement les protéines avec un score SAINT de 0,9 avec un FDR de 0,01 et qui sont donc les plus significativement enrichies dans l'IP. Parmi celles-ci, plusieurs sont déjà connues pour interagir avec la protéine DDB2 tel que le démontre le réseau d'interactions de la figure 19. On retrouve entre autres la protéine DDB1 et les Cullins ce qui démontre une fois de plus l'appartenance de DDB2 à cette famille protéique. De plus l'analyse par *gene ontology* de ses interacteurs révèle des processus biologiques tels que l'organisation des chromosomes ( $p=1.33e-12$ ) et la réparation par excision de nucléotides ( $p=1.75e-12$ ) qui sont les deux fonctions connues de DDB2. Ceci permet de confirmer la validité de notre approche en reproduisant les données de la littérature pour la protéine DDB2.

En plus des protéines déjà connues pour interagir avec DDB2, nous avons identifié quelques protéines qui sont impliquées dans la transcription. Plus particulièrement, les protéines KDM3A et EP300 sont très bien connues pour leur rôle dans ce processus cellulaire via leur activité déméthylase (KDM3A) et acétyltransférase (EP300). (H.-Y. Wang et al. 2019). Ce n'est pas la première fois que DDB2 est rapportée pour interagir avec des protéines impliquées au niveau de la transcription. Une étude publiée à la fin des années 1990 a fait état de l'interaction de la protéine DDB2 avec un facteur de transcription E2F1 suggérant pour la première fois un rôle pour celle-ci dans ce processus biologique (Hayes et al. 1998). Plus d'une dizaine d'années plus tard, une seconde étude a démontré que la protéine DDB2 était

impliquée dans la répression de la transcription d'un gène, via la liaison de DDB2 au promoteur de celui-ci (Minig et al. 2009). En fait, il a été démontré que la protéine DDB2 peut reconnaître des séquences d'ADN spécifiquement au niveau des promoteurs de certains gènes afin de recruter des modificateurs d'histones, tel que des acétyltransférases ou des méthylases. Ce recrutement a comme conséquence la modification des marques d'histones et donc de moduler l'activité transcriptionnelle du gène en question (Zhao et al. 2014). Ainsi, il est probable que la protéine DDB2 peut moduler l'expression de certains gènes via son interaction avec les protéines KDM3A et EP300.

De plus, dans nos résultats de Pulse-SILAC, nous avons identifié la protéine TCERG1 comme ayant une expression différentielle en présence de la surexpression de DDB2 et nous avons confirmé une diminution de la protéine TCERG1 par immunobuvardage de type western. TCERG1 est connue pour son rôle au niveau de la transcription ainsi qu'au niveau de l'épissage alternatif de certains ARNm. Une étude qui s'est intéressée à la réponse transcriptionnelle des cellules à la suite de traitements aux UV a montré que les cellules sans TCERG1 présentaient une réponse transcriptionnelle altérée comparativement à des cellules contrôle, ce qui laisse présager un rôle pour cette protéine dans la réponse cellulaire aux dommages à l'ADN médiée par les UV (Boeing et al. 2016). De plus, l'une des cibles de TCERG1 est le gène de BCL-X, où elle module la vitesse de l'ARN polymérase II, et ainsi influencer l'épissage de BCL-X. La surexpression de TCERG1 promeut la forme courte de la protéine (BCL-X<sub>S</sub> forme pro-apoptotique) et au contraire, sa diminution favorise la forme longue (BCL-X<sub>L</sub> anti-apoptotique) (Montes et al. 2012). L'une des raisons pour lesquelles nous pensons que TCERG1 influencerait BCL-X via la protéine DDB2 est que l'épissage de BCL-X est affecté lors de dommages aux UV favorisant la forme pro-apoptotique de la protéine (BCL-X<sub>S</sub>) (Shkreta et al. 2011). De plus, la protéine DDB2 est connue pour être impliquée au niveau de la répression de la transcription de la protéine BCL-2, une autre protéine anti-apoptotique. En fait, en

utilisant un mutant de DDB2 qui empêche sa liaison à l'ADN et en traitant les cellules avec des agents génotoxiques, il a été démontré que les cellules avec une mutation dans DDB2 présentent une résistance à l'apoptose contrairement aux cellules où la surexpression de DDB2 augmente la mort cellulaire à la suite des traitements (Stoyanova et al. 2009). Donc, si la surexpression de DDB2 diminue la transcription de TCERG1 tout comme elle le fait pour BCL-2, la forme anti-apoptotique de BCL-X (BCL-X<sub>L</sub>) sera favorisée. Au contraire, en absence de DDB2, la protéine TCERG1 sera plus exprimée et la forme pro-apoptotique de BCL-X sera produite (tout comme la protéine BCL-2). L'une des régulations possibles est qu'en temps normal la protéine DDB2 diminue l'expression de la protéine BCL-2 et favorise la forme BCL-X<sub>L</sub> de BCL-X (via la régulation de TCERG1), ce qui envoie des signaux pro-survie à la cellule. Or, lors d'une présence soutenue ou lors d'une accumulation importante de dommage à l'ADN, la protéine DDB2 est fortement recrutée aux sites de dommages, ce qui a comme effet la diminution de la présence de DDB2 au promoteur de BCL-2 et de TCERG1 et favorise les signaux pro-apoptotiques.

Considérant le nombre important de complexes que les protéines DDB1-Cul4 en association avec les DCAFs peuvent former, ainsi que son implication dans de nombreux processus cellulaires importants, il est primordial de bien connaître l'ensemble des fonctions de ce complexe d'E3 ubiquitine-ligase. Dans un premier temps, les résultats obtenus dans ce chapitre nous renseignent sur l'appartenance des protéines à la famille des DCAFs en croisant les données obtenues des IPs avec DDB1 avec celles obtenues dans les IPs DCAFs. De plus, les IPs des protéines DCAFs nous ont permis d'identifier leurs fonctions via l'analyse de leurs partenaires d'interactions, ce qui nous a permis de récapituler les fonctions majoritaires du complexe DDB1-Cul4. Finalement, les expériences de Pulse-SILAC ont rendu possible l'identification des cibles de dégradation des protéines DCAFs. Ensemble, ces travaux ont permis de mieux caractériser non seulement les fonctions mais aussi les cibles du complexe DDB1-Cul4.

## **4.2 Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène**

L'annotation du génome, qui a initialement été faite en prenant en considération 3 principaux facteurs pour l'identification des parties codantes, non codantes et régulatrices, n'a pas été remise en question jusqu'à ce jour. Cependant, la découverte d'ARN régulateurs (tel que les ARNInc) qui peuvent provenir de régions indiquées comme non codantes ou encore de protéines issues des régions régulatrices 5' ou 3' viennent mettre en évidence la nécessité de revoir la façon d'interpréter notre ADN génomique (Andreev et al. 2015). Une autre preuve qui démontre notre manière désuète d'identifier les produits des gènes est la découverte dans le laboratoire du Pr Roucou des protéines alternatives. Ces protéines proviennent en fait d'un autre cadre de lecture (altORF) que celui qui a initialement été identifié (Vanderperre et al. 2011). Ainsi, le nombre de protéines présentes dans la cellule auraient été jusqu'à présent grandement sous-estimé. Afin de répertorier ces protéines, le laboratoire du Pr Roucou a construit une base de données (OpenProt) qui regroupe plusieurs évidences de l'expression de ces protéines avec des études de conservation, de *ribosome profiling* et de spectrométrie de masse (Brunet et al. 2019).

### **4.2.1 Découverte des variants d'ubiquitine produit à partir de UBBP4**

Dans l'optique d'identifier de bons candidats à étudier parmi toutes ces protéines nouvellement répertoriées, nous avons repris une étude d'interactions protéiques à large échelle que nous avons réanalysée pour identifier toutes les protéines alternatives dans les interactomes des protéines de références (Huttlin et al. 2015b). La protéine qui a donné le plus d'interaction avec des protéines de références est une protéine alternative d'une longueur de 77 acides aminés référencée IP\_710742. Les résultats de *Gene Ontology* sur la fonction des protéines de référence ont démontré un enrichissement significatif de différents mécanismes

reliés à l'ubiquitine. En analysant la séquence d'acides aminés de cette alternative, nous avons constaté que celle-ci était pratiquement identique à la protéine d'ubiquitine, hormis pour quatre acides aminés (Q2K, K33E, Q49K et N60S). Mis à part ces différences, la protéine IP\_710742 possède aussi une cystéine additionnelle à la position 77 située après les 2 glycines caractéristiques de ce type de modification post-traductionnelle. Cependant, nous avons testé la présence et l'absence de cet acide aminé sur le profil d'ubiquitinylation de IP\_710742 et celui-ci reste le même en présence ou en absence de la cystéine (résultat non montré). En fait, le gène UBB possède lui aussi une cystéine en dernière position de sa 3<sup>e</sup> répétition qui est clivée, ce qui nous confirme que la cystéine présente sur IP\_710742 n'interfère pas avec son attachement sur d'autres protéines.

IP\_710742 provient en fait du pseudogène d'ubiquitine UBBP4 qui est issu d'une duplication du gène UBB lors de l'évolution. UBBP4 se compose, tout comme UBB, de 3 répétitions de l'ubiquitine avec en plus la protéine IP\_710742 située seulement quelques nucléotides après la 3<sup>e</sup> répétition et localisée dans un second cadre de lecture. Donc, UBBP4 code en fait pour 4 protéines d'ubiquitine dans 2 cadres de lectures différents, nous avons donc renommé chacune de ces répétitions en fonction de leur position et de leur cadre de lecture (A1, A2 et A3 pour le premier ORF et B1 pour la molécule dans le second ORF). UBBP4 a initialement été catégorisée comme un pseudogène en raison de la présence d'une mutation qui mène à un codon-stop à la position 59 dans la répétition A1 (Cowland, Wiborg, et Vuust 1988). Cette mutation a donc mené les scientifiques qui ont découvert ce gène à la classer comme étant un pseudogène. Cependant, nos expériences de séquençage de l'ARNm de UBBP4 ainsi que de son ADN génomique ont démontré que ce codon-stop n'était pas existant et que cette mutation devait en fait être due à une erreur lors de l'analyse de la séquence. La méthode qu'ils ont utilisée à l'époque nécessitait l'interprétation de gel de séquençage qui était un travail complexe et beaucoup moins précis que les méthodes actuelles de séquençage.

De plus, les bases de données qui répertorient les séquences génomiques (tel que NCBI) ont-elles aussi une séquence sans codon-stop pour la molécule A1, mais possèdent plutôt une tyrosine à la position 59. Ces résultats signifient que le gène UBBP4 peut produire les trois molécules d'ubiquitine initialement identifiées en plus de la protéine alternative, UBBP4-B1, que nous avons identifiée et de ce fait, celui-ci ne devrait plus être répertorié comme pseudogène. Une autre mutation qui avait été rapportée dans le gène UBBP4 est l'absence d'un acide aminé dans une répétition normale comptant 7 cystéines dans la protéine UBBP4-A3, ce qui entraîne donc un décalage du cadre de lecture et la production d'une protéine sans les deux glycines en C-terminal. Puisque les répétitions de nucléotides sont souvent sujettes à des erreurs lors du séquençage et que visiblement les techniques de l'époque étaient moins sensibles, nous avons pensé que peut-être cette absence était aussi une aberration dans la séquence. Cependant, nos expériences ont démontré qu'il y avait bel et bien que 6 cystéines et que la protéine UBBP4-A3 ne possédait pas une séquence qui ressemble à l'ubiquitine en 3'.

Bien que les sous-unités A1, A2 et B1 produites par UBBP4 possèdent des caractéristiques similaires à l'ubiquitine, telles qu'une longueur de 76 acides aminés et la présence des deux glycines en C-terminal, ces dernières se différencient de la protéine parentale par la présence de mutations dans leur séquence en acides aminés. En fait, la protéine A1 possède 8 acides aminés différents de l'ubiquitine qui sont dispersés tout au long de sa séquence. La protéine A2 est celle qui ressemble le plus à la canonique avec seulement un acide aminé qui diffère et la protéine B1 telle que mentionnée précédemment en possède 4. À ce jour, UBBP4 est le seul gène qui produit des molécules d'ubiquitine avec une séquence en acides aminés différentes de celles qui sont encodées par les gènes connus.

Une autre preuve que UBBP4 est différent des autres pseudogènes de cette famille est que lors de la comparaison des niveaux d'ARNm provenant des gènes UBBP1 à UBBP5, on remarque que le nombre de transcrits identifiés pour UBBP4 est

nettement supérieur comparativement aux autres membres de la famille. Cette différence ne peut être attribué à l'absence de signaux uniques pour l'ARNm des gènes de UBBP1,2,3, et 5 puisque chacun de ceux-ci, tout comme UBBP4, comporte des différences dans leurs séquences comparativement à l'ubiquitine qui permet de les différencier. Ainsi, l'ARNm de UBBP4 est beaucoup plus exprimé comparativement aux autres pseudogènes sans toutefois atteindre les mêmes niveaux que l'ubiquitine.

Nous avons aussi déterminé que non seulement l'ARNm de UBBP4 était présent (par l'utilisation des bases de données, mais aussi lors de nos expériences de séquençage), mais que cet ARNm était recouvert par des ribosomes d'élongations. En effet, les données de ribosome profiling récupérées de plusieurs expériences démontrent un bon signal pour la section du gène qui comporte la protéine A1, A3 et B1 avec une diminution, voir une absence de signal pour tout le long de la protéine A2 (Michel et al. 2014). Cette absence de signal peut être due au fait que les protéines A2 possèdent seulement 1 acide aminé de différence avec l'ubiquitine canonique. Ainsi, il est probable que les séquences de l'ARNm qui correspondent à A2 s'apparient sur les autres gènes de l'ubiquitine ce qui amène une absence de signal pour cette portion de l'ADN. Ainsi les résultats de *ribosome profiling* analysés indiquent que l'ARNm provenant du gène UBBP4 est occupé sur toute sa longueur par des ribosomes. Finalement, en récupérant les données de spectrométrie de masse disponibles sur OpenProt, nous avons identifié plusieurs peptides provenant des sous-unités A1, A2 et B1 qui ont été détectées dans de nombreuses études. La détection de ces peptides uniques est possible grâce à la présence des différences dans leurs séquences, ce qui les différencie des peptides provenant de l'ubiquitine canonique. De manière surprenante, aucun peptide pour la molécule A3 n'a été détecté. Bien que le début de la séquence soit identique avec l'ubiquitine, ce qui empêche l'identification de peptide unique, le décalage de lecture entraîne la production d'une séquence totalement différente entre les deux protéines, ce qui devrait permettre l'identification de peptides correspondant à cette

section de la protéine. De plus, l'utilisation d'outils bio-informatiques ont démontré que cette portion de la protéine A3 possède des peptides détectables par spectrométrie de masse, ce qui ne peut expliquer le fait qu'on n'identifie pas de peptides pour cette section. L'une des raisons qui pourrait expliquer ce résultat est que les peptides générés ne sont pas uniques à cette protéine ce qui rend impossible leur identification dans la base de données OpenProt. Néanmoins, les peptides qui sont identifiés pour les sous-unités A1, A2 et B1 nous confirment que le gène UBBP4 est bel et bien exprimé sous forme de protéines dans la cellule.

L'une des approches les plus couramment utilisées et la mieux acceptée par la communauté scientifique pour prouver la présence d'une protéine dans la cellule reste l'utilisation d'anticorps spécifiques dans des techniques biochimiques classiques telles que l'immunobuvardage et l'immunofluorescence. L'utilisation de cette méthode pour prouver l'expression des sous-unités du gène UBBP4 est très laborieuse puisque les protéines ressemblent énormément à l'ubiquitine canonique. Nous avons cependant tenté l'expérience avec la protéine UBBP4-B1 en produisant un anticorps dirigé contre un peptide contenant l'un des quatre acides aminés différents. Même si les résultats sur les protéines recombinantes étaient prometteurs (résultats non montrés), l'utilisation de l'anticorps dans des extraits cellulaires s'est avérée moins concluante. D'un côté, parce que les modifications post-traductionnelles peuvent s'attacher sur plusieurs protéines dans la cellule, ce qui donne par immunobuvardage, un profil qui ressemble plus à un extrait cellulaire qu'à une hybridation spécifique. Ainsi, en l'absence de poids moléculaire de "référence" pour notre protéine, il était difficile de conclure sur la spécificité de notre anticorps. Une autre façon de prouver l'existence des protéines produites par UBBP4 aurait été de regarder les marques protéiques laissées par celle-ci sur les peptides en spectrométrie de masse, à la suite de la digestion par une enzyme telle que la trypsine. Cependant, encore une fois les marques laissées par l'ubiquitine (GG) sont les mêmes que celles laissées par les protéines encodées par UBBP4, ce qui nous empêche, par cette approche, de démontrer leur expression. Une approche qu'on



pourrait utiliser serait d'ajouter une étiquette sur chacune des protéines de UBBP4 endogène à l'aide de la technologie CRISPR et d'utiliser des anticorps contre celles-ci pour monitorer leur expression. Malheureusement, les séquences trop similaires entre les protéines et les autres gènes de l'ubiquitine empêchent ou du moins complexifient grandement cette approche. Toutefois, les données d'expression d'ARNm, de ribosome profiling et de spectrométrie de masse nous apportent des preuves probantes de la transcription de UBBP4 et de sa traduction en protéine.

Comme mentionné, l'ubiquitine, ainsi que les protéines ubiquitine-like sont capables de s'attacher de manière covalente à une multitude de protéines dans la cellule. Cet attachement est possible grâce à la présence des deux glycines en C-terminal de la protéine qui sont nécessaires pour la formation du lien. Tout comme les protéines ubiquitine-like, les protéines A2 et B1 sont capables de s'attacher à de nombreuses protéines dans la cellule tel que le démontre l'immunobuvardage à la figure 25. Les protéines A1 et A3 sont pour leur part présentes seulement sous forme de monomères, ce qui était attendu pour la protéine A3 puisque celle-ci ne contient pas les deux glycines à la fin de sa séquence protéique. Toutefois, l'absence de profil de type ubiquitine pour la protéine A1 est surprenante puisque celle-ci contient les acides aminés essentiels pour la formation de lien covalent. D'un point de vue biochimique, les 8 mutations présentes dans la séquence d'acides aminés de cette protéine pourrait impliquer un changement dans sa structure tertiaire, ce qui pourrait l'empêcher d'être reconnue par la machinerie enzymatique responsable de son activation et de son attachement. Il est aussi possible que ces mutations, sans changer la structure, interfèrent avec les liaisons des enzymes E1, E2 et E3. Il a été démontré pour l'ubiquitine que certaines mutations ponctuelles sont capables de diminuer ou d'empêcher complètement la liaison de cette dernière avec l'enzyme E1 (Singh et al. 2017). Parmi ces acides aminés impliqués dans l'interaction, certains se retrouvent mutés dans la séquence de A1, ce qui pourrait avoir comme effet de diminuer ou d'empêcher l'activation de la modification post-traductionnelle.

#### 4.2.2 Réponse des variants A2 et B1 à l'inhibition du protéasome

Bien que l'ubiquitine puisse moduler certaines caractéristiques biochimiques de ses substrats, telles que la localisation cellulaire ou les interactions protéiques, la majorité des événements d'ubiquitylation vont mener à une altération de la stabilité par la dégradation protéasomale. Il est donc normal que l'inhibition de ce processus (via l'utilisation d'inhibiteur spécifique du protéasome) se traduise par une augmentation du signal d'ubiquitylation lorsque monitoré par immunobuvardage. Il semble que la mutation présente dans la séquence de la protéine A2 n'altère pas cette fonction puisque le profil est semblable à celui de l'ubiquitine à la suite du traitement au MG132, un inhibiteur du protéasome. La protéine B1 présente pour sa part, un profil complètement différent de celui de l'ubiquitine avec une diminution de l'intensité du signal en immunobuvardage après traitement. De plus, l'absence d'agrégats en immunofluorescence et les résultats de spectrométrie de masse ont confirmé que les protéines modifiées par B1 n'adressent pas les protéines au protéasome. La protéine B1 diffère de l'ubiquitine par seulement 4 acides aminés, ce qui signifie que ces mutations sont capables d'induire une signalisation cellulaire différente de l'ubiquitine. L'analyse de la structure de B1 n'a démontré aucune différence significative entre les deux protéines (travaux réalisés dans le laboratoire du Pr Lavigne), ce qui ne pouvait expliquer la différence de signalisation. En fait, nous pensons que la protéine B1, que nous avons renommée Ub<sup>KEKS</sup>, interfère dans cette fonction cellulaire par la formation de nouvelles chaînes. Puisque celle-ci possède deux nouvelles lysines (K2 et K49), nous pensons que la formation de chaînes sur la lysine 49 interférerait avec la liaison au protéasome (voisine de la lysine 48, lien majoritaire dans cette fonction cellulaire). Ces nouveaux liens agiraient comme un mécanisme de protection contre la dégradation des protéines. Lors de l'inhibition du protéasome, ces liens de "protection" qui sont faits via Ub<sup>KEKS</sup> se font peu à peu remplacer par des molécules d'ubiquitine pour envoyer les protéines à la dégradation, ce qui pourrait expliquer la diminution du signal produit par la molécule d'Ub<sup>KEKS</sup>. Cette théorie est renforcée par le fait que les protéines

marquées par Ub<sup>KEKS</sup> ne forment pas d'agrégats en présence de MG132, ce qui signifie que ces protéines ne sont pas destinées à la dégradation puisque si c'était le cas, on verrait tout de même la formation de ces structures via l'ubiquitine.

Une autre possibilité est que la molécule Ub<sup>KEKS</sup> engendre une signalisation complètement différente de l'ubiquitine, ce qui pourrait expliquer l'absence d'accumulation de signal et d'agrégats lors de l'inhibition du protéasome. Toutefois les ressemblances entre l'ubiquitine et Ub<sup>KEKS</sup> sont tellement importantes qu'il est difficile d'envisager une régulation complètement séparée pour les deux protéines.

#### **4.2.3 PCNA, FANCI et FANCD2**

L'identification des protéines modifiées par Ub<sup>KEKS</sup> avait comme objectif de nous renseigner sur les possibles fonctions de la modification post-traductionnelle. Dans nos expériences de spectrométrie de masse (aussi bien en conditions dénaturantes qu'en conditions natives), la protéine qui est toujours la plus enrichie dans les deux lignées cellulaires testées est la protéine PCNA. PCNA est bien connue pour son rôle dans la réplication de l'ADN où elle augmente la processivité de la polymérase et où elle sert aussi de plateforme de recrutement pour une multitude de protéines (Celis et al. 1987). La régulation de la protéine PCNA a été intimement liée aux différentes PTM qui prennent place sur la protéine et plus particulièrement sur les lysines de cette dernière. En fait, des analyses de spectrométrie de masse ont démontré que sur toutes les lysines présentes sur la protéine, 13 d'entre elles (sur une possibilité de 16) peuvent être soit ubiquitinylées, SUMOylées, acétylées, NEDDylées, méthylées ou ISGylées. Bien que toutes ces PTM n'aient pas été identifiées sur les 13 lysines, certains sites sont la cible de plusieurs de celles-ci (Kanao et Masutani 2017). Par exemple, la lysine 248 peut être modifiée par 5 de ces PTM qui vont réguler de manière différentielle les interactions de PCNA, le niveau d'expression ou encore sa présence à la chromatine (Choe et Moldovan 2017; Takawa et al. 2012). Ainsi, la modification par PTM sur un seul site sur PCNA est capable d'engendrer différentes réponses cellulaires. C'est ce principe qui est utilisé

par la protéine PCNA lors de la réponse aux dommages à l'ADN. En fait, PCNA est bien connue pour son rôle dans le choix des mécanismes de réparations suite aux blocages de la fourche de réplication et c'est l'ubiquitinylation sur sa lysine 164 qui oriente cette décision (Kanao et Masutani 2017). De plus, non seulement le site de modification est déterminant, mais la présence d'une mono- ou d'une poly-ubiquitine sur celle-ci engage différents mécanismes de réparation. L'identification de PCNA comme cible majoritaire de Ub<sup>KEKS</sup> dans notre étude remet en question la régulation de PCNA par l'ubiquitine. En fait, plusieurs études démontrent la modification de PCNA par l'ubiquitine seulement avec la présence d'une bande de plus hauts poids moléculaires en réponse à différents dommages. De plus, puisque nous avons identifié que le site de modification par Ub<sup>KEKS</sup> se situe sur une lysine en C-terminal qui comprend entre autres la lysine 164 et que les cellules présentent des défauts de croissance cellulaire, nous pensons que cette dernière pourrait être impliquée dans la régulation des mécanismes de réparation de l'ADN. Il a été démontré que la lysine 164 peut aussi être SUMOylée chez l'humain pour empêcher la recombinaison homologue lors du blocage de la réplication via l'interaction de protéines spécifiques (Moldovan et al. 2012). Ainsi, tout comme la lysine 248, la lysine 164 peut être modifiée par plusieurs PTM, ce qui nous laisse croire que Ub<sup>KEKS</sup> pourrait aussi prendre place sur cette lysine. Bien que nous ne remettions pas en question l'entièreté de la régulation apportée par l'ubiquitine sur PCNA, il est probable que certaines fonctions qui ont été attribuées à cette modification viennent plutôt d'Ub<sup>KEKS</sup>.

De plus, dans les cellules sans Ub<sup>KEKS</sup>, nous avons identifié une diminution de la signature GlyGly sur les protéines FANCI et FANCD2, deux protéines impliquées dans la réponse aux dommages à l'ADN. En outre, les sites que nous avons identifiés sur ces deux protéines sont ceux qui sont connus pour être ubiquitinylés en réponse aux dommages afin d'amorcer la réparation de l'ADN par le mécanisme de Fanconie (Ishiai et al. 2017).

Malheureusement, comme mentionné précédemment, la signature laissée sur les peptides par l'ubiquitine et par Ub<sup>KEKS</sup> est la même, ce qui nous empêche de conclure sur l'identité de la modification qui est directement impliquée sur ces protéines. Néanmoins, le fait que nous n'ayons jamais identifié ces protéines dans les IPs avec Ub<sup>KEKS</sup> nous laisse croire que la diminution du signal est attribuable à l'ubiquitine et que c'est par une régulation indirecte de Ub<sup>KEKS</sup> que ces deux sites diminuent.

Les résultats que nous avons obtenus ne nous permettent cependant pas de conclure si Ub<sup>KEKS</sup> régulerait la réponse aux dommages à l'ADN ou encore si la régulation apportée par celle-ci était plus au niveau de la réplication de l'ADN. En effet, un défaut au niveau des mécanismes de réplication amènerait les mêmes phénotypes que nous avons observés entre autres pour un retard de la croissance. Cependant, la protéine PCNA a été démontrée pour être peu modifiée lors de la réplication de l'ADN, ce qui nous laisse croire à un rôle lors du blocage de la fourche de réplication. De plus, si la régulation prenait place lors de la réplication, il serait plus logique que la cellule utilise l'ubiquitine, qui est probablement plus abondante, pour un mécanisme fréquent, plutôt que Ub<sup>KEKS</sup> qui est possiblement en petite quantité dans la cellule.

#### **4.2.4 Lamine et Ub<sup>KEKS</sup>**

Une autre protéine qui a attiré notre attention lors de l'identification des cibles de Ub<sup>KEKS</sup> est la protéine LMNB2. D'une part parce que non seulement celle-ci était enrichie de manière significative dans nos IPs, mais aussi que deux autres lamines (LMNA et LMNB1) étaient aussi présentes dans nos résultats avec des taux d'enrichissement légèrement plus faibles. Une autre raison qui nous a porté à s'intéresser de plus près à cette famille de protéines provient de l'observation par immunofluorescences de Ub<sup>KEKS</sup> qui se retrouve enrichie à la membrane nucléaire, qui est la localisation des protéines de la famille des lamines. Afin de déterminer si Ub<sup>KEKS</sup> pouvait avoir un rôle dans la régulation des lamines, nous avons réalisé des

immunofluorescences sur ces protéines dans nos cellules défectées pour ce gène et nous avons démontré que LMNA se retrouve non seulement à la membrane nucléaire, mais aussi qu'une sous-population de la protéine se localise dans les nucléoles. Nous avons essayé de confirmer ce changement de localisation par des expériences de fractionnements, mais après plusieurs essais, ils nous étaient impossible d'enrichir une fraction nucléolaire des cellules. Ceci pourrait traduire une certaine "fragilité" des nucléoles dans ces cellules. Par ailleurs, les nucléoles dans ces cellules sont significativement plus gros comparativement aux cellules normales. Ces phénotypes sont d'ailleurs tous présents dans les laminopathies, qui sont une série de désordres génétiques touchant le gène LMNA qui se manifestent par la dégénérescence tissulaire ou la sénescence prématurée des tissus (Luo, Mastaglia, et Wilton 2014). Dans ces maladies, les mutations interfèrent avec le bon fonctionnement de la protéine LMNA entre autres par un problème au niveau de sa maturation, puisque celle-ci doit subir plusieurs modifications post-traductionnelles et un clivage protéolytique pour être mature (Hennekes et Nigg 1994). Le mauvais fonctionnement de la protéine se traduit par une présence persistante de dommages à l'ADN, un défaut de réparation, une perte de l'hétérochromatine et une structure du noyau aberrante (Gonzalo 2014). Rappelons que Ub<sup>KEKS</sup> est capable de cibler des protéines de dommages à l'ADN tel que PCNA et que les cellules déplétées pour cette modification présentent des retards de croissance, signe associé à un problème des mécanismes de réparation de l'ADN. De plus, l'une des formes de laminopathies (Hutchinson-Gilford progeria syndrome) présente des nucléoles anormalement larges comparativement à des cellules normales, phénotype que nous retrouvons aussi dans nos cellules UBBP4<sup>-/-</sup>. Ce phénotype serait non seulement le résultat d'une augmentation du métabolisme de ces cellules, mais aussi par une localisation aberrante de LMNA dans le nucléoplasme (Buchwalter et Hetzer 2017).

Les modifications post-transcriptionnelles de LMNA par l'ubiquitine n'ont pas été largement étudiées. Il a cependant été démontré que l'Ub interfère dans la stabilité de la protéine. En fait, des souris présentant une délétion d'une des lysines

(K32) présentent des cardiomyopathies sévères qui se produisent par la dégradation prématurée de LMNA mutée comparativement à la protéine de type sauvage (Cattin et al. 2013). De plus, lors du développement, les protéines LMNA sont présentes dans le nucléoplasme et se relocalisent à la membrane lors de la formation des filaments. Dans les mutants  $\Delta K32$ , la protéine reste de façon beaucoup plus marquée à l'intérieur du noyau (Bertrand et al. 2012). Ce dernier phénotype rappelle celui observé dans les cellules sans Ub<sup>KEKS</sup> où la protéine LMNA, qui est de type sauvage dans les cellules, se relocalise aux nucléoles. Ainsi, on pourrait postuler que Ub<sup>KEKS</sup> cible LMNA sur cet acide aminé pour réguler l'un des aspects de la protéine (maturation, localisation, formation des filaments) et que la délétion de celle-ci empêche le bon fonctionnement de LMNA, ce qui se traduit par des phénotypes similaires aux laminopathies. Bien que nous n'ayons pas clairement démontré que l'abolition de Ub<sup>KEKS</sup> sur LMNA soit directement reliée à ce type de maladies, les phénotypes présents dans les cellules sans UBBP4<sup>-/-</sup> sont très similaires à ceux des laminopathies.

En somme, ces travaux ont permis la découverte de deux nouveaux variants fonctionnels de l'ubiquitine qui proviennent d'un gène orthologue d'UBB et qui possèdent des différences en acides aminées comparativement à la molécule d'ubiquitine canonique. De plus, nos expériences ont démontré que le variant Ub<sup>KEKS</sup> n'adresse pas les protéines à la dégradation par le protéasome et que sa délétion ralentit la croissance cellulaire. Nous avons identifié que la protéine PCNA peut être modifiée par Ub<sup>KEKS</sup> et ce sur les mêmes lysines que l'ubiquitine, ce qui met en perspective l'importance de différencier la présence de l'ubiquitine ou de Ub<sup>KEKS</sup> sur les protéines cibles. Finalement, nous avons découvert que les lamines sont elles aussi les cibles de Ub<sup>KEKS</sup> et que la localisation de la protéine LMNA est modifiée dans les cellules sans le gène UBBP4. Tous ces résultats mettent l'accent sur l'importance de cette nouvelle modification post-traductionnelle.

## 5. Conclusion

### **Chapitre 1 : Étude des cibles du complexe DDB1-CulA via l'intermédiaire des protéines DCAFs**

Les complexes d'ubiquitines ligases sont le cœur de la régulation amenée par l'ubiquitine par la reconnaissance des différentes cibles de cette modification. Considérant le rôle majeur de l'ubiquitine dans presque tous les processus cellulaires, il n'est pas étonnant de constater qu'un très grand nombre de E3-ligases sont l'objet des recherches dans le but d'identifier leurs fonctions spécifiques. Le premier chapitre de cette thèse visait l'étude des fonctions du complexe E3-ubiquitine formé de DDB1-Cul4 par l'intermédiaire des protéines responsables de la reconnaissance des substrats, les protéines DCAFs. Parmi les protéines DCAFs proposées dans la littérature, nous avons identifié dans nos expériences de spectrométrie de masse 27 de ces protéines qui interagissent avec le complexe CRL4 en condition physiologique. Comme mentionné précédemment, il n'est cependant pas à exclure que d'autres protéines, dans la liste initiale puissent interagir avec le complexe dans d'autres conditions (ex : en présence de dommages à l'ADN). Afin de les identifier, il serait intéressant de traiter les cellules avec différents agents et analyser dans des IPs DDB1 afin de déterminer si on peut en identifier d'autres. De plus, d'autres protéines ont été identifiées via nos expériences de spectrométrie de masse pour interagir avec le complexe, dont certaines avec un domaine WD40. Il pourrait être très intéressant de cloner ces protéines et refaire le même type d'expérience effectué dans cette thèse pour confirmer leur appartenance à cette famille.

On dénombre 33 processus cellulaires qui ont été identifiés via l'analyse des protéines qui interagissent avec les protéines DCAFs dans nos expériences. Ces fonctions peuvent se regrouper sous 10 grandes catégories qui ont été pour la plupart déjà associées au complexe DDB1-Cul4. Nous avons cependant mis en



évidence une implication de certaines DCAFs dans les mécanismes d'épissage alternatif telle que la protéine DCAF4 qui n'avait pas encore de fonction associée. Afin de conclure sur la véritable implication du complexe DDB1 dans ces fonctions, nous allons devoir faire des études plus spécifiques à chacune des DCAFs tel que confirmer l'ubiquitinylation de leurs substrats. Une approche qui a été utilisée précédemment et qui semble bien fonctionner pour l'identification des cibles est de diminuer l'expression de la protéine (ici les DCAFs) et d'enrichir par différentes méthodes les protéines et les peptides ubiquitinylés avant de les analyser par spectrométrie de masse. L'identification directe des motifs d'ubiquitine sur les protéines et les peptides (double confirmation) pourrait directement nous indiquer si la modification de celles-ci est diminuée en absence de la DCAF. En ce qui concerne les IPs pour les protéines qui ne sont finalement pas associées à la protéine DDB1, les résultats de ces expériences sont très intéressants et méritent tout de même d'être investigués. En particulier les protéines qui ont été identifiées pour être essentielles à la survie cellulaire ou celles qui n'ont pas encore de rôle cellulaire défini.

Pour la protéine DDB2, nous avons mis de l'avant son rôle dans la régulation de la protéine TCERG1 impliquée dans la transcription et dans l'épissage alternatif des ARNm. On ne peut cependant aucunement déterminer si cette régulation est via son activité ubiquitine-ligase ou via la régulation de la transcription de la protéine. Afin de trancher par quel mécanisme la protéine TCERG1 est régulée par DDB2, il faudrait regarder l'ARNm de TCERG1 lors de la surexpression de DDB2 ou encore, regarder si une inhibition du protéasome sauve le phénotype de disparition de la protéine. Ainsi nous pourrions conclure par laquelle des activités de DDB2 celle-ci est régulée. Finalement, bien que l'hypothèse de la protéine BCL-X semble plausible, il faudrait quantifier la formation des deux isoformes de la protéine pour conclure sur son implication via TCERG1 et les dommages aux UVs. Somme toute, cette découverte vient mettre en lumière un nouveau rôle de la protéine DDB2 dans la régulation de la protéine TCERG1 et potentiellement dans l'apoptose.

En conclusion, ce travail nous a permis de confirmer l'association des protéines DCAFs avec les protéines DDB-Cul4 dans des conditions normales. De plus, nous avons identifié de nouvelles fonctions pour le complexe DDB1-Cul4 via la caractérisation de rôle cellulaire de protéines qui n'avait pas été étudié auparavant. Ce travail ouvre une multitude de possibilités de projets qu'ils seraient très intéressants de continuer en particulier celui sur la protéine DDB2 et TCERG1.

## **Chapitre 2 : Identification d'un nouveau variant d'ubiquitine exprimé à partir d'un pseudogène**

En ce qui concerne le chapitre 2, nous avons découvert que non seulement UBBP4 est en fait un gène, mais que celui-ci peut produire 4 molécules d'ubiquitine dans deux cadres de lecture différents. Deux de ces molécules sont d'ailleurs capables de s'attacher à une multitude de protéines dans la cellule. Plus particulièrement la sous-unité que nous avons renommée Ub<sup>KEKS</sup> n'adresse pas les protéines à la dégradation via le protéasome. Nous n'avons cependant aucune indication du mécanisme que cette modification emploie pour ne pas cibler le protéasome. L'une des potentielle hypothèse est que la présence de nouvelles lysines sur Ub<sup>KEKS</sup> permet la formation de chaînes qui interfèrent avec la fixation au protéasome. Pour prouver cette hypothèse, il faudrait démontrer que Ub<sup>KEKS</sup> peut être modifiée sur les lysines K48 et K49 en même temps et que cette liaison empêche le récepteur du protéasome de lier ce branchement. Néanmoins, l'expression de ces variants vient remettre en question l'analyse que nous devons faire par rapport à certains outils que nous utilisons pour conclure sur l'ubiquitinylation d'une protéine. En fait, les anticorps qui reconnaissent l'ubiquitine peuvent aussi reconnaître Ub<sup>KEKS</sup>. De plus, la signature protéique laissée par ces deux modifications est identique lorsque monitorée par spectrométrie de masse. Ces deux derniers points mettent en lumière que nous avons probablement par le passé, attribué des fonctions cellulaires à l'ubiquitine qui pourrait être en fait causée par Ub<sup>KEKS</sup>.

En ce qui concerne la protéine PCNA, bien que nous connaissions le rôle de l'ubiquitine sur celle-ci et que nous avons identifié que Ub<sup>KEKS</sup> la modifie dans le même domaine, nous ne pouvons conclure sur une fonction précise de cette nouvelle PTM. Cependant, le fait que plusieurs histones soient aussi identifiées dans les IPs de Ub<sup>KEKS</sup> et que les cellules présentent un retard de croissance sont des indications que Ub<sup>KEKS</sup> pourrait réguler certains aspects de la protéine PCNA, soit dans la réplication ou dans la réparation de l'ADN. Pour déterminer dans quelle voie cette modification est impliquée nous pourrions vérifier les dommages à l'ADN en réponse à différents traitements dans les lignées délétées pour UB<sup>KEKS</sup> ou encore regarder la progression de la fourche de réplication dans ces cellules par des essais de fibre d'ADN (*DNA fiber assay*). Pour vérifier si Ub<sup>KEKS</sup> modifie les partenaires d'interaction de PCNA, nous pourrions les étudier par spectrométrie de masse en condition normale ou en présence de dommages dans les cellules UBBP4<sup>-/-</sup>. Le rôle central de PCNA dans le choix des mécanismes de réparation et que celui-ci soit régulé par l'ubiquitine vient souligner de manière importante la nécessité de bien faire la distinction entre l'ubiquitine et Ub<sup>KEKS</sup> et le rôle potentiellement important de cette nouvelle modification post-traductionnelle.

Pour les lamines, nous avons démontré que plusieurs des membres de cette famille pouvaient être modifiés par Ub<sup>KEKS</sup>. De plus, la protéine LMNA est relocalisée dans les nucléoles dans les cellules délétées pour UBBP4 en plus que ces cellules présentent des nucléoles anormalement plus gros. Ces phénotypes se comparent à ceux présents dans les laminopathies, désordres génétiques touchant la protéine LMNA. Comme mentionné, LMNA est important pour le maintien de l'hétérochromatine, il pourrait donc être intéressant d'aller monitorer celle-ci dans les cellules délétées pour Ub<sup>KEKS</sup> afin de voir l'état de celle-ci. L'identification du site de modification par Ub<sup>KEKS</sup> sur la protéine LMNA pourrait nous renseigner si un mutant de ce ou ces acides aminés ont été détectés dans les laminopathies, ce qui nous permettrait de faire un lien entre nos phénotypes observés et ceux de ces

maladies en plus d'impliquer UBBP4 dans une pathologie déjà bien connue. Bien qu'il nous soit encore malheureusement impossible de déterminer le rôle de cette modification sur les lamines, celle-ci semble importante pour le bon fonctionnement de cette famille protéique.

En conclusion pour ce chapitre, celui-ci présente des nouveaux variants de l'ubiquitine qui sont produits à partir du gène d'UBBP4. Cette découverte majeure vient remettre plusieurs choses en question, en particulier quant aux méthodes d'études des modifications post-traductionnelles. En fait, il n'est pas impossible que UBBP4 ne soit pas une exception et qu'il existe d'autres pseudogènes de l'ubiquitine ou d'autres ubiquitines-like qui sont exprimées et qui peuvent avoir un rôle au sein de la cellule. De plus, il se pourrait que la machinerie cellulaire associée à l'ubiquitine (E1, E2 ou E3) présente en fait des préférences pour l'ubiquitine canonique ou pour les différents variants de UBBP4. Finalement, il pourrait aussi être très intéressant de quantifier UB<sup>KEKS</sup> par rapport à l'ubiquitine dans la cellule de manière endogène afin d'évaluer l'importance de cette nouvelle modification comparativement à la canonique. Somme toute, cette découverte ouvre une multitude de questions et d'avenues quant aux modifications post-traductionnelles et le rôle d'UBBP4 dans la cellule.

Les résultats obtenus dans les deux chapitres, même s'ils ne sont pas directement reliés, s'influencent l'un l'autre. En fait, la découverte de nouveaux variants d'ubiquitine vient remettre en question une partie des connaissances concernant la régulation de cette modification post-traductionnelle et de son utilisation par les complexes de E3 tel que DDB1-Cul4. Il est probable qu'on découvre dans un avenir rapproché, qu'un site d'ubiquitination qui a été attribué à DDB1-Cul4 soit en fait une modification apportée par Ub<sup>KEKS</sup>. Dans cette éventualité une meilleure connaissance des cibles de ce complexe de E3 sera déjà un bon point de départ pour commencer une nouvelle étude.

## **6. Bibliographie**

- Aichem, Annette, et Marcus Groettrup. 2016. « The Ubiquitin-like Modifier FAT10 in Cancer Development ». *The International Journal of Biochemistry & Cell Biology* 79: 451-61.
- Akutsu, Masato, Ivan Dikic, et Anja Bremm. 2016. « Ubiquitin Chain Diversity at a Glance ». *Journal of Cell Science* 129(5): 875-80.
- Al-Hakim, Abdallah K. et al. 2008. « Control of AMPK-Related Kinases by USP9X and Atypical Lys(29)/Lys(33)-Linked Polyubiquitin Chains ». *The Biochemical Journal* 411(2): 249-60.
- An, Jian et al. 2017. « PSILAC Mass Spectrometry Reveals ZFP91 as IMiD-Dependent Substrate of the CRL4<sup>CRBN</sup> Ubiquitin Ligase ». *Nature Communications* 8: 15398.
- Andreev, Dmitry E. et al. 2015. « Translation of 5' Leaders Is Pervasive in Genes Resistant to EIF2 Repression ». *eLife* 4: e03971.
- Angers, Stephane et al. 2006. « Molecular Architecture and Assembly of the DDB1-CUL4A Ubiquitin Ligase Machinery ». *Nature* 443(7111): 590-93.
- Baker, Rohan T., et Philip G. Board. 1991. « The Human Ubiquitin-52 Amino Acid Fusion Protein Gene Shares Several Structural Features with Mammalian Ribosomal Protein Genes ». *Nucleic Acids Research* 19(5): 1035-40.
- Barbier-Torres, Lucía et al. 2015. « Stabilization of LKB1 and Akt by Neddylation Regulates Energy Metabolism in Liver Cancer ». *Oncotarget* 6(4): 2509-23.
- Beal, R. et al. 1996. « Surface Hydrophobic Residues of Multiubiquitin Chains Essential for Proteolytic Targeting ». *Proceedings of the National Academy of Sciences of the United States of America* 93(2): 861-66.
- Bellare, Priya et al. 2006. « Ubiquitin Binding by a Variant Jab1/MPN Domain in the Essential Pre-mRNA Splicing Factor Prp8p ». *RNA (New York, N.Y.)* 12(2): 292-302.
- Bellare, Priya et al. 2008. « A Role for Ubiquitin in the Spliceosome Assembly Pathway ». *Nature Structural & Molecular Biology* 15(5): 444-51.
- Bertrand, Anne T. et al. 2012. « DelK32-Lamin A/C Has Abnormal Location and Induces Incomplete Tissue Maturation and Severe Metabolic Defects Leading to Premature Death ». *Human Molecular Genetics* 21(5): 1037-48.
- Bianchi, Marzia et al. 2015. « Dynamic Transcription of Ubiquitin Genes under Basal and Stressful Conditions and New Insights into the Multiple UBC Transcript Variants ». *Gene* 573(1): 100-109.

- Bindea, Gabriela et al. 2009. « ClueGO: A Cytoscape Plug-in to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks ». *Bioinformatics (Oxford, England)* 25(8): 1091-93.
- Boeing, Stefan et al. 2016. « Multiomic Analysis of the UV-Induced DNA Damage Response ». *Cell Reports* 15(7): 1597-1610.
- Boname, Jessica M. et al. 2010. « Efficient Internalization of MHC I Requires Lysine-11 and Lysine-63 Mixed Linkage Polyubiquitin Chains ». *Traffic (Copenhagen, Denmark)* 11(2): 210-20.
- Brunet, Marie A. et al. 2019. « OpenProt: A More Comprehensive Guide to Explore Eukaryotic Coding Potential and Proteomes ». *Nucleic Acids Research* 47(D1): D403-10.
- Buchwalter, Abigail, et Martin W. Hetzer. 2017. « Nucleolar Expansion and Elevated Protein Translation in Premature Aging ». *Nature Communications* 8(1): 1-13.
- Cajee, Umar-Faruq, Rodney Hull, et Monde Ntwasa. 2012. « Modification by Ubiquitin-Like Proteins: Significance in Apoptosis and Autophagy Pathways ». *International Journal of Molecular Sciences* 13(9): 11804.
- Cang, Yong et al. 2006. « Deletion of DDB1 in Mouse Brain and Lens Leads to P53-Dependent Elimination of Proliferating Cells ». *Cell* 127(5): 929-40.
- Cappadocia, Laurent, et Christopher D. Lima. 2018. « Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism ». *Chemical Reviews* 118(3): 889-918.
- Cattin, Marie-Elodie et al. 2013. « Heterozygous LmnadelK32 Mice Develop Dilated Cardiomyopathy through a Combined Pathomechanism of Haploinsufficiency and Peptide Toxicity ». *Human Molecular Genetics* 22(15): 3152-64.
- Celis, J. E. et al. 1987. « Cyclin (PCNA, Auxiliary Protein of DNA Polymerase Delta) Is a Central Component of the Pathway(s) Leading to DNA Replication and Cell Division ». *FEBS letters* 220(1): 1-7.
- Chiu, Yu-Hsin, Qinmiao Sun, et Zhijian J. Chen. 2007. « E1-L2 Activates Both Ubiquitin and FAT10 ». *Molecular Cell* 27(6): 1014-23.
- Choe, Katherine N., et George-Lucian Moldovan. 2017. « Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork ». *Molecular Cell* 65(3): 380-92.
- Choi, Hyungwon et al. 2011. « SAINT: Probabilistic Scoring of Affinity Purification–Mass Spectrometry Data ». *Nature Methods* 8(1): 70-73.

- Ciechanover, A. et al. 1980. « Characterization of the Heat-Stable Polypeptide of the ATP-Dependent Proteolytic System from Reticulocytes ». *The Journal of Biological Chemistry* 255(16): 7525-28.
- Ciechanover, A., Y. Hod, et A. Hershko. 1978. « A Heat-Stable Polypeptide Component of an ATP-Dependent Proteolytic System from Reticulocytes ». *Biochemical and Biophysical Research Communications* 81(4): 1100-1105.
- Cowland, J. B., O. Wiborg, et J. Vuust. 1988. « Human Ubiquitin Genes: One Member of the UbB Gene Subfamily Is a Tetrameric Non-Processed Pseudogene ». *FEBS letters* 231(1): 187-91.
- Damgaard, Rune Busk et al. 2012. « The Ubiquitin Ligase XIAP Recruits LUBAC for NOD2 Signaling in Inflammation and Innate Immunity ». *Molecular Cell* 46(6): 746-58.
- Dammer, Eric B. et al. 2011. « Polyubiquitin Linkage Profiles in Three Models of Proteolytic Stress Suggest the Etiology of Alzheimer Disease ». *The Journal of Biological Chemistry* 286(12): 10457-65.
- D'Cunha, J. et al. 1996. « Immunoregulatory Properties of ISG15, an Interferon-Induced Cytokine ». *Proceedings of the National Academy of Sciences of the United States of America* 93(1): 211-15.
- Dias, Dora C., Georgia Dolios, Rong Wang, et Zhen-Qiang Pan. 2002. « CUL7: A DOC Domain-Containing Cullin Selectively Binds Skp1.Fbx29 to Form an SCF-like Complex ». *Proceedings of the National Academy of Sciences of the United States of America* 99(26): 16601-6.
- Dikic, Ivan, Soichi Wakatsuki, et Kylie J. Walters. 2009. « Ubiquitin-Binding Domains — from Structures to Functions ». *Nature Reviews Molecular Cell Biology* 10(10): 659-71.
- Dornan, David et al. 2004. « The Ubiquitin Ligase COP1 Is a Critical Negative Regulator of P53 ». *Nature* 429(6987): 86-92.
- Dornan, David et al. 2006. « ATM Engages Autodegradation of the E3 Ubiquitin Ligase COP1 After DNA Damage ». *Science* 313(5790): 1122-26.
- Drissi, Romain, Marie-Line Dubois, Mélanie Douziech, et François-Michel Boisvert. 2015. « Quantitative Proteomics Reveals Dynamic Interactions of the Minichromosome Maintenance Complex (MCM) in the Cellular Response to Etoposide Induced DNA Damage ». *Molecular & Cellular Proteomics : MCP* 14(7): 2002-13.
- Dubois, Marie-Line, Charlotte Bastin, Dominique Lévesque, et François-Michel Boisvert. 2016. « Comprehensive Characterization of Minichromosome Maintenance Complex (MCM) Protein Interactions Using Affinity and Proximity Purifications Coupled to Mass Spectrometry ». *Journal of Proteome Research* 15(9): 2924-34.

- Duda, David M. et al. 2008. « Structural Insights into NEDD8 Activation of Cullin-RING Ligases: Conformational Control of Conjugation ». *Cell* 134(6): 995-1006.
- Durcan, Thomas M. et al. 2014. « USP8 Regulates Mitophagy by Removing K6-Linked Ubiquitin Conjugates from Parkin ». *The EMBO journal* 33(21): 2473-91.
- Dzimianski, John V., Florine E. M. Scholte, Éric Bergeron, et Scott D. Pegan. 2019. « ISG15: It's Complicated ». *Journal of Molecular Biology* 431(21): 4203-16.
- van Eersel, Janet et al. 2011. « Cytoplasmic Accumulation and Aggregation of TDP-43 upon Proteasome Inhibition in Cultured Neurons ». *PLoS ONE* 6(7).  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3146516/> (9 août 2019).
- Elia, Andrew E.H. et al. 2015. « Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response ». *Molecular cell* 59(5): 867-81.
- Embade, Nieves et al. 2012. « Murine Double Minute 2 Regulates Hu Antigen R Stability in Human Liver and Colon Cancer through NEDDylation ». *Hepatology (Baltimore, Md.)* 55(4): 1237-48.
- Emmerich, Christoph H. et al. 2013a. « Activation of the Canonical IKK Complex by K63/M1-Linked Hybrid Ubiquitin Chains ». *Proceedings of the National Academy of Sciences of the United States of America* 110(38): 15247-52.
- Enchev, Radoslav I., Brenda A. Schulman, et Matthias Peter. 2015. « Protein Neddylation: Beyond Cullin-RING Ligases ». *Nature Reviews. Molecular Cell Biology* 16(1): 30-44.
- Falquet, L. et al. 1995. « A Human De-Ubiquitinating Enzyme with Both Isopeptidase and Peptidase Activities in Vitro ». *FEBS letters* 359(1): 73-77.
- Farrell, P. J., R. J. Broeze, et P. Lengyel. 1979. « Accumulation of an mRNA and Protein in Interferon-Treated Ehrlich Ascites Tumour Cells ». *Nature* 279(5713): 523-25.
- Flick, Karin et al. 2006. « A Ubiquitin-Interacting Motif Protects Polyubiquitinated Met4 from Degradation by the 26S Proteasome ». *Nature Cell Biology* 8(5): 509-15.
- Fornace, A J, I Alamo, M C Hollander, et E Lamoreaux. 1989. « Ubiquitin mRNA is a major stress-induced transcript in mammalian cells. » *Nucleic Acids Research* 17(3): 1215-30.
- Fu, DongTao et al. 2003. « CDNA Cloning of the Chicken DDB1 Gene Encoding the P127 Subunit of Damaged DNA-Binding Protein ». *Genes & Genetic Systems* 78(2): 169-77.
- Furukawa, K., N. Mizushima, T. Noda, et Y. Ohsumi. 2000. « A Protein Conjugation System in Yeast with Homology to Biosynthetic Enzyme Reaction of Prokaryotes ». *The Journal of Biological Chemistry* 275(11): 7462-65.



- Galisson, Frederic et al. 2011. « A Novel Proteomics Approach to Identify SUMOylated Proteins and Their Modification Sites in Human Cells ». *Molecular & cellular proteomics: MCP* 10(2): M110.004796.
- Gatti, Marco et al. 2015. « RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage ». *Cell Reports* 10(2): 226-38.
- Gerlach, Björn et al. 2011. « Linear Ubiquitination Prevents Inflammation and Regulates Immune Signalling ». *Nature* 471(7340): 591-96.
- Glutzer, M., A. W. Murray, et M. W. Kirschner. 1991. « Cyclin Is Degraded by the Ubiquitin Pathway ». *Nature* 349(6305): 132-38.
- Goldenberg, Seth J. et al. 2004. « Structure of the Cand1-Cul1-Roc1 Complex Reveals Regulatory Mechanisms for the Assembly of the Multisubunit Cullin-Dependent Ubiquitin Ligases ». *Cell* 119(4): 517-28.
- Gonzalo, Susana. 2014. « DNA Damage and Lamins ». *Advances in experimental medicine and biology* 773: 377-99.
- Grice, Guinevere L. et al. 2015. « The Proteasome Distinguishes between Heterotypic and Homotypic Lysine-11-Linked Polyubiquitin Chains ». *Cell Reports* 12(4): 545-53.
- Grou, Cláudia P. et al. 2015. « The *de Novo* Synthesis of Ubiquitin: Identification of Deubiquitinases Acting on Ubiquitin Precursors ». *Scientific Reports* 5: 12836.
- Haas, A. L., J. V. Warms, A. Hershko, et I. A. Rose. 1982. « Ubiquitin-Activating Enzyme. Mechanism and Role in Protein-Ubiquitin Conjugation ». *The Journal of Biological Chemistry* 257(5): 2543-48.
- Haas, Tobias L. et al. 2009. « Recruitment of the Linear Ubiquitin Chain Assembly Complex Stabilizes the TNF-R1 Signaling Complex and Is Required for TNF-Mediated Gene Induction ». *Molecular Cell* 36(5): 831-44.
- Hanada, Takao et al. 2007. « The Atg12-Atg5 Conjugate Has a Novel E3-like Activity for Protein Lipidation in Autophagy ». *The Journal of Biological Chemistry* 282(52): 37298-302.
- Hannah, Jeffrey, et Pengbo Zhou. 2015. « Distinct and overlapping functions of the cullin E3 ligase scaffolding proteins CUL4A and CUL4B ». *Gene* 573(1): 33-45.
- Hart, Traver et al. 2015. « High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities ». *Cell* 163(6): 1515-26.
- Hayes, S., P. Shiyanov, X. Chen, et P. Raychaudhuri. 1998. « DDB, a Putative DNA Repair Protein, Can Function as a Transcriptional Partner of E2F1 ». *Molecular and Cellular Biology* 18(1): 240-49.

- He, Yizhou Joseph et al. 2006. « DDB1 Functions as a Linker to Recruit Receptor WD40 Proteins to CUL4–ROC1 Ubiquitin Ligases ». *Genes & Development* 20(21): 2949-54.
- Hendriks, Ivo A. et al. 2014. « Uncovering Global SUMOylation Signaling Networks in a Site-Specific Manner ». *Nature Structural & Molecular Biology* 21(10): 927-36.
- Hennekes, H., et E. A. Nigg. 1994. « The Role of Isoprenylation in Membrane Attachment of Nuclear Lamins. A Single Point Mutation Prevents Proteolytic Cleavage of the Lamin A Precursor and Confers Membrane Binding Properties ». *Journal of Cell Science* 107 ( Pt 4): 1019-29.
- Herhaus, Lina, et Ivan Dikic. 2015. « Expanding the ubiquitin code through post-translational modification ». *EMBO Reports* 16(9): 1071-83.
- Hershko, A. et al. 1980. « Proposed Role of ATP in Protein Breakdown: Conjugation of Protein with Multiple Chains of the Polypeptide of ATP-Dependent Proteolysis ». *Proceedings of the National Academy of Sciences of the United States of America* 77(4): 1783-86.
- Hipp, Mark Steffen et al. 2005. « FAT10, a Ubiquitin-Independent Signal for Proteasomal Degradation ». *Molecular and Cellular Biology* 25(9): 3483-91.
- Hu, Jian, Chad M. McCall, Tomohiko Ohta, et Yue Xiong. 2004. « Targeted Ubiquitination of CDT1 by the DDB1-CUL4A-ROC1 Ligase in Response to DNA Damage ». *Nature Cell Biology* 6(10): 1003-9.
- Huang, Danny T. et al. 2009. « E2-RING Expansion of the NEDD8 Cascade Confers Specificity to Cullin Modification ». *Molecular Cell* 33(4): 483-95.
- Humbard, Matthew A. et al. 2010. « Ubiquitin-like Small Archaeal Modifier Proteins (SAMPs) in *Haloferax Volcanii* ». *Nature* 463(7277): 54-60.
- Husnjak, Koraljka, et Ivan Dikic. 2012. « Ubiquitin-Binding Proteins: Decoders of Ubiquitin-Mediated Cellular Functions ». *Annual Review of Biochemistry* 81: 291-322.
- Huttlin, Edward L. et al. 2015a. « The BioPlex Network: A Systematic Exploration of the Human Interactome ». *Cell* 162(2): 425-40.
- Ikeda, Fumiyo et al. 2011. « SHARPIN Forms a Linear Ubiquitin Ligase Complex Regulating NF-KB Activity and Apoptosis ». *Nature* 471(7340): 637-41.
- Ingolia, Nicholas T. 2016. « Ribosome Footprint Profiling of Translation throughout the Genome ». *Cell* 165(1): 22-33.
- Ishiai, Masamichi et al. 2017. « Activation of the FA Pathway Mediated by Phosphorylation and Ubiquitination ». *Mutation Research* 803-805: 89-95.

- Ishimi, Yukio. 2018. « Regulation of MCM2-7 Function ». *Genes & Genetic Systems* 93(4): 125-33.
- Jacobson, Andrew D. et al. 2009. « The Lysine 48 and Lysine 63 Ubiquitin Conjugates Are Processed Differently by the 26 S Proteasome ». *Journal of Biological Chemistry* 284(51): 35485-94.
- Jin, Jianping et al. 2006. « A Family of Diverse Cul4-Ddb1-Interacting Proteins Includes Cdt2, Which Is Required for S Phase Destruction of the Replication Factor Cdt1 ». *Molecular Cell* 23(5): 709-21.
- Jin, Jianping, Xue Li, Steven P. Gygi, et J. Wade Harper. 2007. « Dual E1 Activation Systems for Ubiquitin Differentially Regulate E2 Enzyme Charging ». *Nature* 447(7148): 1135-38.
- Jin, Lingyan et al. 2008a. « Mechanism of Ubiquitin-Chain Formation by the Human Anaphase-Promoting Complex ». *Cell* 133(4): 653-65.
- Johnson, E. S., P. C. Ma, I. M. Ota, et A. Varshavsky. 1995. « A Proteolytic Pathway That Recognizes Ubiquitin as a Degradation Signal ». *The Journal of Biological Chemistry* 270(29): 17442-56.
- Johnson, R. T. et al. 1985. « Xeroderma Pigmentosum D-HeLa Hybrids with Low and High Ultraviolet Sensitivity Associated with Normal and Diminished DNA Repair Ability, Respectively ». *Journal of Cell Science* 76: 115-33.
- Jurica, Melissa S. et al. 2002. « Purification and Characterization of Native Spliceosomes Suitable for Three-Dimensional Structural Analysis ». *RNA (New York, N.Y.)* 8(4): 426-39.
- Ka, Won Hye et al. 2018. « The Ubiquitin Ligase COP1 Regulates Cell Cycle and Apoptosis by Affecting P53 Function in Human Breast Cancer Cell Lines ». *Breast Cancer (Tokyo, Japan)* 25(5): 529-38.
- Kanao, Rie, et Chikahide Masutani. 2017. « Regulation of DNA Damage Tolerance in Mammalian Cells by Post-Translational Modifications of PCNA ». *Mutation Research* 803-805: 82-88.
- Kim, Hyoungh Tae et al. 2007. « Certain Pairs of Ubiquitin-Conjugating Enzymes (E2s) and Ubiquitin-Protein Ligases (E3s) Synthesize Nondegradable Forked Ubiquitin Chains Containing All Possible Isopeptide Linkages ». *The Journal of Biological Chemistry* 282(24): 17375-86.
- Kirisako, Takayoshi et al. 1999. « Formation Process of Autophagosome Is Traced with Apg8/Aut7p in Yeast ». *The Journal of Cell Biology* 147(2): 435-46.
- Kirisako, Takayoshi et al. 2006. « A ubiquitin ligase complex assembles linear polyubiquitin chains ». *The EMBO Journal* 25(20): 4877-87.

- Kirkpatrick, Donald S. et al. 2006. « Quantitative Analysis of in Vitro Ubiquitinated Cyclin B1 Reveals Complex Chain Topology ». *Nature Cell Biology* 8(7): 700-710.
- Kisselev, A. F., et A. L. Goldberg. 2001. « Proteasome Inhibitors: From Research Tools to Drug Candidates ». *Chemistry & Biology* 8(8): 739-58.
- Komatsu, Masaaki et al. 2004. « A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier ». *The EMBO Journal* 23(9): 1977-86.
- Koyano, Fumika et al. 2014. « Ubiquitin Is Phosphorylated by PINK1 to Activate Parkin ». *Nature* 510(7503): 162-66.
- Kriegenburg, Franziska, Christian Ungermann, et Fulvio Reggiori. 2018. « Coordination of Autophagosome-Lysosome Fusion by Atg8 Family Members ». *Current biology: CB* 28(8): R512-18.
- Lake, M. W., M. M. Wuebbens, K. V. Rajagopalan, et H. Schindelin. 2001. « Mechanism of Ubiquitin Activation Revealed by the Structure of a Bacterial MoeB-MoaD Complex ». *Nature* 414(6861): 325-29.
- Lamoliatte, Frederic et al. 2013. « Targeted Identification of SUMOylation Sites in Human Proteins Using Affinity Enrichment and Paralog-Specific Reporter Ions ». *Molecular & cellular proteomics: MCP* 12(9): 2536-50.
- Larsen, C. N., B. A. Krantz, et K. D. Wilkinson. 1998. « Substrate Specificity of Deubiquitinating Enzymes: Ubiquitin C-Terminal Hydrolases ». *Biochemistry* 37(10): 3358-68.
- Lee, Imsang, et Hermann Schindelin. 2008. « Structural Insights into E1-Catalyzed Ubiquitin Activation and Transfer to Conjugating Enzymes ». *Cell* 134(2): 268-78.
- Lee, Jae-Hoon et al. 2008. « Characterization of Arabidopsis and Rice DWD Proteins and Their Roles as Substrate Receptors for CUL4-RING E3 Ubiquitin Ligases ». *The Plant Cell* 20(1): 152-67.
- Lee, Jennifer, et Pengbo Zhou. 2007a. « DCAFs, the Missing Link of the CUL4-DDB1 Ubiquitin Ligase ». *Molecular Cell* 26(6): 775-80.
- Leimkühler, S., M. M. Wuebbens, et K. V. Rajagopalan. 2001. « Characterization of Escherichia Coli MoeB and Its Involvement in the Activation of Molybdopterin Synthase for the Biosynthesis of the Molybdenum Cofactor ». *The Journal of Biological Chemistry* 276(37): 34695-701.
- Li, Wei et al. 2008a. « Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 That Regulates the Organelle's Dynamics and Signaling ». *PloS One* 3(1): e1487.

- Li, Zheng, et Xingzhi Xu. 2019. « Post-Translational Modifications of the Mini-Chromosome Maintenance Proteins in DNA Replication ». *Genes* 10(5).
- Liakopoulos, D, G Doenges, K Matuschewski, et S Jentsch. 1998. « A novel protein modification pathway related to the ubiquitin system. » *The EMBO Journal* 17(8): 2208-14.
- Lingaraju, Gondichatnahalli M. et al. 2014. « Crystal Structure of the Human COP9 Signalosome ». *Nature* 512(7513): 161-65.
- Liu, Hongliang et al. 2017. « Functional variants in DCAF4 associated with lung cancer risk in European populations ». *Carcinogenesis* 38(5): 541-51.
- Liu, Jidong, Manabu Furukawa, Tomohiro Matsumoto, et Yue Xiong. 2002. « NEDD8 Modification of CUL1 Dissociates P120(CAND1), an Inhibitor of CUL1-SKP1 Binding and SCF Ligases ». *Molecular Cell* 10(6): 1511-18.
- Liu, Xianpeng et al. 2017. « Orthogonal ubiquitin transfer identifies ubiquitination substrates under differential control by the two ubiquitin activating enzymes ». *Nature Communications* 8.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5290280/> (17 juillet 2019).
- Liu, Xiuxia et al. 2016. « The Ubiquitin-like Protein FAT10 Stabilizes EEF1A1 Expression to Promote Tumor Proliferation in a Complex Manner ». *Cancer Research* 76(16): 4897-4907.
- Liu, Zhengzhao et al. 2014. « Ubiquitylation of Autophagy Receptor Optineurin by HACE1 Activates Selective Autophagy for Tumor Suppression ». *Cancer Cell* 26(1): 106-20.
- Loeb, K. R., et A. L. Haas. 1992. « The Interferon-Inducible 15-KDa Ubiquitin Homolog Conjugates to Intracellular Proteins ». *The Journal of Biological Chemistry* 267(11): 7806-13.
- Lonsdale, John et al. 2013. « The Genotype-Tissue Expression (GTEx) Project ». *Nature Genetics* 45: 580-85.
- Lu, Ying et al. 2015. « Substrate Degradation by the Proteasome: A Single-Molecule Kinetic Analysis ». *Science (New York, N.Y.)* 348(6231): 1250834.
- Luo, Yue-Bei, Frank L. Mastaglia, et Steve D. Wilton. 2014. « Normal and Aberrant Splicing of LMNA ». *Journal of Medical Genetics* 51(4): 215-23.
- Lv, Zongyang et al. 2018. « Crystal Structure of a Human Ubiquitin E1-Ubiquitin Complex Reveals Conserved Functional Elements Essential for Activity ». *Journal of Biological Chemistry*: jbc.RA118.003975.

- Ly, Valentina et al. 2015. « Light and COP1 regulate level of overexpressed DET1 protein ». *Plant Science* 231: 114-23.
- Machida, Yuichi J. et al. 2006. « UBE2T Is the E2 in the Fanconi Anemia Pathway and Undergoes Negative Autoregulation ». *Molecular Cell* 23(4): 589-96.
- Malakhova, Oxana, Michael Malakhov, Christopher Hetherington, et Dong-Er Zhang. 2002. « Lipopolysaccharide Activates the Expression of ISG15-Specific Protease UBP43 via Interferon Regulatory Factor 3 ». *The Journal of Biological Chemistry* 277(17): 14703-11.
- Matunis, M. J., E. Coutavas, et G. Blobel. 1996. « A Novel Ubiquitin-like Modification Modulates the Partitioning of the Ran-GTPase-Activating Protein RanGAP1 between the Cytosol and the Nuclear Pore Complex ». *The Journal of Cell Biology* 135(6 Pt 1): 1457-70.
- McDowell, Gary S., et Anna Philpott. 2013. « Non-canonical ubiquitylation: Mechanisms and consequences ». *The International Journal of Biochemistry & Cell Biology* 45(8): 1833-42.
- Meyer, Hermann-Josef, et Michael Rape. 2011. « Processive Ubiquitin Chain Formation by the Anaphase-Promoting Complex ». *Seminars in cell & developmental biology* 22(6): 544-50.
- eyer, Hermann-Josef, et Michael Rape. 2014. « Enhanced protein degradation by branched ubiquitin chains ». *Cell* 157(4): 910-21.
- Michel, Audrey M. et al. 2014. « GWIPS-Viz: Development of a Ribo-Seq Genome Browser ». *Nucleic Acids Research* 42(Database issue): D859-864.
- Michelle, Caroline, Patrick Vourc'h, Laurence Mignon, et Christian R. Andres. 2009. « What Was the Set of Ubiquitin and Ubiquitin-like Conjugating Enzymes in the Eukaryote Common Ancestor? ». *Journal of Molecular Evolution* 68(6): 616-28.
- Min, Mingwei et al. 2015. « Efficient APC/C substrate degradation in cells undergoing mitotic exit depends on K11 ubiquitin linkages ». *Molecular Biology of the Cell* 26(24): 4325-32.
- Minig, Vanessa et al. 2009. « Identification of DDB2 Protein as a Transcriptional Regulator of Constitutive SOD2 Gene Expression in Human Breast Cancer Cells ». *The Journal of Biological Chemistry* 284(21): 14165-76.
- Mizushima, N. et al. 1998. « A Protein Conjugation System Essential for Autophagy ». *Nature* 395(6700): 395-98.
- Moldovan, George-Lucian et al. 2012. « Inhibition of Homologous Recombination by the PCNA-Interacting Protein PARI ». *Molecular Cell* 45(1): 75-86.

- Montes, Marta et al. 2012. « TCERG1 Regulates Alternative Splicing of the Bcl-x Gene by Modulating the Rate of RNA Polymerase II Transcription ». *Molecular and Cellular Biology* 32(4): 751-62.
- Morris, Joanna R., et Ellen Solomon. 2004. « BRCA1 : BARD1 Induces the Formation of Conjugated Ubiquitin Structures, Dependent on K6 of Ubiquitin, in Cells during DNA Replication and Repair ». *Human Molecular Genetics* 13(8): 807-17.
- Murrow, Lyndsay, et Jayanta Debnath. 2015. « ATG12-ATG3 Connects Basal Autophagy and Late Endosome Function ». *Autophagy* 11(6): 961-62.
- Nakasone, Mark A. et al. 2013. « Mixed-Linkage Ubiquitin Chains Send Mixed Messages ». *Structure (London, England: 1993)* 21(5): 727-40.
- Nucifora Jr, Frederick C. et al. 2016. « Ubiquitination via K27 and K29 Chains Signals Aggregation and Neuronal Protection of LRRK2 by WSB1 ». *Nature Communications* 7: 11792.
- Nunoura, Takuro et al. 2011. « Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group ». *Nucleic Acids Research* 39(8): 3204-23.
- Ohi, Melanie D. et al. 2003. « Structural insights into the U-box, a domain associated with multi-ubiquitination ». *Nature structural biology* 10(4): 250-55.
- Ohno, Yoshinori et al. 2010. « Hoxb4 Transduction Down-Regulates Geminin Protein, Providing Hematopoietic Stem and Progenitor Cells with Proliferation Potential ». *Proceedings of the National Academy of Sciences* 107(50): 21529-34.
- Ohta, Tomohiko, Jennifer J. Michel, Arndt J. Schottelius, et Yue Xiong. 1999. « ROC1, a Homolog of APC11, Represents a Family of Cullin Partners with an Associated Ubiquitin Ligase Activity ». *Molecular Cell* 3(4): 535-41.
- Ohtake, Fumiaki et al. 2015. « Ubiquitin Acetylation Inhibits Polyubiquitin Chain Elongation ». *EMBO reports* 16(2): 192-201.
- Ohtake, Fumiaki, Hikaru Tsuchiya, Yasushi Saeki, et Keiji Tanaka. 2018. « K63 Ubiquitylation Triggers Proteasomal Degradation by Seeding Branched Ubiquitin Chains ». *Proceedings of the National Academy of Sciences* 115(7): E1401-8.
- Ordureau, Alban et al. 2015. « Defining Roles of PARKIN and Ubiquitin Phosphorylation by PINK1 in Mitochondrial Quality Control Using a Ubiquitin Replacement Strategy ». *Proceedings of the National Academy of Sciences of the United States of America* 112(21): 6637-42.
- Oved, Shlomo et al. 2006. « Conjugation to Nedd8 Instigates Ubiquitylation and Down-Regulation of Activated Receptor Tyrosine Kinases ». *The Journal of Biological Chemistry* 281(31): 21640-51.

- Park, Hyejin, Min-Sik Yoon, et Kwon-Yul Ryu. 2013. « Disruption of Polyubiquitin Gene Ubc Leads to Defective Proliferation of Hepatocytes and Bipotent Fetal Liver Epithelial Progenitor Cells ». *Biochemical and Biophysical Research Communications* 435(3): 434-40.
- Patterson, M., et G. Chu. 1989. « Evidence That Xeroderma Pigmentosum Cells from Complementation Group E Are Deficient in a Homolog of Yeast Photolyase ». *Molecular and Cellular Biology* 9(11): 5105-12.
- Pauli, Andrea et al. 2014. « Toddler: An Embryonic Signal That Promotes Cell Movement via Apelin Receptors ». *Science (New York, N.Y.)* 343(6172): 1248636.
- Petroski, Matthew D., et Raymond J. Deshaies. 2005a. « Function and Regulation of Cullin-RING Ubiquitin Ligases ». *Nature Reviews. Molecular Cell Biology* 6(1): 9-20.
- Petroski, Matthew D., et Raymond J. Deshaies. 2005c. « Mechanism of Lysine 48-Linked Ubiquitin-Chain Synthesis by the Cullin-RING Ubiquitin-Ligase Complex SCF-Cdc34 ». *Cell* 123(6): 1107-20.
- Petryszak, Robert et al. 2016. « Expression Atlas Update—an Integrated Database of Gene and Protein Expression in Humans, Animals and Plants ». *Nucleic Acids Research* 44(D1): D746-52.
- Pichler, Andrea, Chronis Fatouros, Heekyoung Lee, et Nathalie Eisenhardt. 2017. « SUMO conjugation – a mechanistic view ». *Biomolecular Concepts* 8(1): 13–36.
- Pierce, Nathan W. et al. 2013. « Cand1 Promotes Assembly of New SCF Complexes through Dynamic Exchange of F Box Proteins ». *Cell* 153(1): 206-15.
- Precious, B. et al. 2005. « Simian Virus 5 V Protein Acts as an Adaptor, Linking DDB1 to STAT2, To Facilitate the Ubiquitination of STAT1 ». *Journal of Virology* 79(21): 13434-41.
- Raasi, S., G. Schmidtke, R. de Giuli, et M. Groettrup. 1999. « A Ubiquitin-like Protein Which Is Synergistically Inducible by Interferon-Gamma and Tumor Necrosis Factor-Alpha ». *European Journal of Immunology* 29(12): 4030-36.
- Radoshevich, Lilliana et al. 2010. « ATG12 Conjugation to ATG3 Regulates Mitochondrial Homeostasis and Cell Death ». *Cell* 142(4): 590-600.
- Ran, F. Ann et al. 2013. « Genome Engineering Using the CRISPR-Cas9 System ». *Nature Protocols* 8(11): 2281-2308.
- Ravichandran, Ramya, Kiran Kodali, Junmin Peng, et Patrick Ryan Potts. 2019. « Regulation of MAGE-A3/6 by the CRL4-DCAF12 ubiquitin ligase and nutrient availability ». *EMBO reports* 20(7): e47352.



- Ravid, Tommer, et Mark Hochstrasser. 2007. « Autoregulation of an E2 Enzyme by Ubiquitin-Chain Assembly on Its Catalytic Residue ». *Nature Cell Biology* 9(4): 422-27.
- Reyes-Turcu, Francisca E. et al. 2006. « The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin ». *Cell* 124(6): 1197-1208.
- Rubinstein, Assaf D. et al. 2011. « The Autophagy Protein Atg12 Associates with Antiapoptotic Bcl-2 Family Members to Promote Mitochondrial Apoptosis ». *Molecular Cell* 44(5): 698-709.
- Ryu, Kwon-Yul et al. 2007. « The Mouse Polyubiquitin Gene Ubc Is Essential for Fetal Liver Development, Cell-Cycle Progression and Stress Tolerance ». *The EMBO journal* 26(11): 2693-2706.
- Ryu, Kwon-Yul et al. 2008. « Hypothalamic Neurodegeneration and Adult-Onset Obesity in Mice Lacking the Ubb Polyubiquitin Gene ». *Proceedings of the National Academy of Sciences of the United States of America* 105(10): 4016-21.
- Ryu, K.-Y. et al. 2010. « Loss of Polyubiquitin Gene Ubb Leads to Metabolic and Sleep Abnormalities in Mice ». *Neuropathology and Applied Neurobiology* 36(4): 285-99.
- Sarikas, Antonio, Thomas Hartmann, et Zhen-Qiang Pan. 2011. « The Cullin Protein Family ». *Genome Biology* 12(4): 220.
- Schabla, N. Max, Koushik Mondal, et Patrick C. Swanson. 2018. « DCAF1 (VprBP): Emerging Physiological Roles for a Unique Dual-Service E3 Ubiquitin Ligase Substrate Receptor ». *Journal of Molecular Cell Biology*.
- Schlieker, Christian D. et al. 2008. « A functional proteomics approach links the ubiquitin-related modifier Urm1 to a tRNA modification pathway ». *Proceedings of the National Academy of Sciences of the United States of America* 105(47): 18255-60.
- Schulman, Brenda A., et J. Wade Harper. 2009. « Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways ». *Nature reviews. Molecular cell biology* 10(5): 319-31.
- Scrima, Andrea et al. 2008. « Structural Basis of UV DNA-Damage Recognition by the DDB1-DDB2 Complex ». *Cell* 135(7): 1213-23.
- Senga, Takeshi et al. 2006. « PCNA Is a Cofactor for Cdt1 Degradation by CUL4/DDB1-Mediated N-Terminal Ubiquitination ». *The Journal of Biological Chemistry* 281(10): 6246-52.
- Seo, Jinho et al. 2018. « K6 Linked Polyubiquitylation of FADD by CHIP Prevents Death Inducing Signaling Complex Formation Suppressing Cell Death ». *Oncogene* 37(36): 4994-5006.

- Shabek, Nitzan et al. 2018. « Structural Insights into DDA1 Function as a Core Component of the CRL4-DDB1 Ubiquitin Ligase ». *Cell Discovery* 4(1): 1-4.
- Sheng, Caibin et al. 2019. « PCNA-Mediated Degradation of P21 Coordinates the DNA Damage Response and Cell Cycle Regulation in Individual Cells ». *Cell Reports* 27(1): 48-58.e7.
- Shevchenko, A. et al. 1996. « Linking Genome and Proteome by Mass Spectrometry: Large-Scale Identification of Yeast Proteins from Two Dimensional Gels ». *Proceedings of the National Academy of Sciences of the United States of America* 93(25): 14440-45.
- Shiyanov, Pavel, Alo Nag, et Pradip Raychaudhuri. 1999. « Cullin 4A Associates with the UV-Damaged DNA-Binding Protein DDB ». *Journal of Biological Chemistry* 274(50): 35309-12.
- Shkreta, Lulzim et al. 2011. « The DNA Damage Response Pathway Regulates the Alternative Splicing of the Apoptotic Mediator Bcl-x ». *The Journal of Biological Chemistry* 286(1): 331-40.
- Singh, Rajesh K, Yaniv Kazansky, Donald Wathieu, et David Fushman. 2017. « The Hydrophobic Patch of Ubiquitin is Important for its Optimal Activation by Ubiquitin Activating Enzyme E1 ». *Analytical chemistry* 89(15): 7852-60.
- Sloper-Mould, Katherine E., Jennifer C. Jemc, Cecile M. Pickart, et Linda Hicke. 2001. « Distinct Functional Surface Regions on Ubiquitin ». *Journal of Biological Chemistry* 276(32): 30483-89.
- Stoyanova, Tanya et al. 2009. « DDB2 Decides Cell Fate Following DNA Damage ». *Proceedings of the National Academy of Sciences of the United States of America* 106(26): 10690-95.
- Su, Andrew I. et al. 2004. « A gene atlas of the mouse and human protein-encoding transcriptomes ». *Proceedings of the National Academy of Sciences of the United States of America* 101(16): 6062-67.
- Sun, Y., M. Tan, H. Duan, et M. Swaroop. 2001. « SAG/ROC/Rbx/Hrt, a Zinc RING Finger Gene Family: Molecular Cloning, Biochemical Properties, and Biological Functions ». *Antioxidants & Redox Signaling* 3(4): 635-50.
- Takawa, Masashi et al. 2012. « Histone Lysine Methyltransferase SETD8 Promotes Carcinogenesis by Deregulating PCNA Expression ». *Cancer Research* 72(13): 3217-27.
- Tang, Jean, et Gilbert Chu. 2002. « Xeroderma Pigmentosum Complementation Group E and UV-Damaged DNA-Binding Protein ». *DNA repair* 1(8): 601-16.

- Tatham, Michael H. et al. 2001. « Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9 ». *Journal of Biological Chemistry* 276(38): 35368-74.
- Taylor, S. V. et al. 1998. « Thiamin Biosynthesis in Escherichia Coli. Identification of ThiS Thiocarboxylate as the Immediate Sulfur Donor in the Thiazole Formation ». *The Journal of Biological Chemistry* 273(26): 16555-60.
- Uhlén, Mathias et al. 2015. « Proteomics. Tissue-Based Map of the Human Proteome ». *Science (New York, N.Y.)* 347(6220): 1260419.
- Van der Veen, Annemarie G. et al. 2011. « Role of the ubiquitin-like protein Urm1 as a noncanonical lysine-directed protein modifier ». *Proceedings of the National Academy of Sciences of the United States of America* 108(5): 1763-70.
- Vanderperre, Benoît et al. 2011. « An Overlapping Reading Frame in the PRNP Gene Encodes a Novel Polypeptide Distinct from the Prion Protein ». *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 25(7): 2373-86.
- Walczak, Christopher P. et al. 2019. « Ribosomal Protein RPL26 Is the Principal Target of UFMylation ». *Proceedings of the National Academy of Sciences* 116(4): 1299-1308.
- Wang, Hui-Yi et al. 2019. « Histone demethylase KDM3A is required for enhancer activation of hippo target genes in colorectal cancer ». *Nucleic Acids Research* 47(5): 2349-64.
- Wang, Xiaodan et al. 2015. « Both Rbx1 and Rbx2 exhibit a functional role in the HIV-1 Vif-Cullin5 E3 ligase complex in vitro ». *Biochemical and Biophysical Research Communications* 461(4): 624-29.
- Wei, Ying, et Xingzhi Xu. 2016. « UFMylation: A Unique & Fashionable Modification for Life ». *Genomics, Proteomics & Bioinformatics* 14(3): 140-46.
- Wertz, Ingrid E. et al. 2015. « Phosphorylation and Linear Ubiquitin Direct A20 Inhibition of Inflammation ». *Nature* 528(7582): 370-75.
- Wilkinson, Keith D. 2004. « Ubiquitin: A Nobel Protein ». *Cell* 119(6): 741-45.
- Wilkinson, Kevin A., et Jeremy M. Henley. 2010. « Mechanisms, Regulation and Consequences of Protein SUMOylation ». *The Biochemical Journal* 428(2): 133-45.
- Will, Cindy L., et Reinhard Lührmann. 2011. « Spliceosome Structure and Function ». *Cold Spring Harbor Perspectives in Biology* 3(7).  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3119917/> (24 septembre 2019).

- Wingender, E, P Dietze, H Karas, et R Knüppel. 1996. « TRANSFAC: a database on transcription factors and their DNA binding sites. » *Nucleic Acids Research* 24(1): 238-41.
- Wittschieben, Birgitte Ø., et Richard D. Wood. 2003. « DDB complexities ». *DNA Repair* 2(9): 1065-69.
- Xirodimas, Dimitris P. et al. 2008. « Ribosomal Proteins Are Targets for the NEDD8 Pathway ». *EMBO reports* 9(3): 280-86.
- Xu, Ping et al. 2009. « Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation ». *Cell* 137(1): 133-45.
- Yang, Chun-Song et al. 2017. « Ubiquitin Modification by the E3 Ligase/ADP-ribosyltransferase Dtx3L/Parp9 ». *Molecular cell* 66(4): 503-516.e5.
- Yang, Xiaofeng et al. 2013. « Absolute Quantification of E1, Ubiquitin-like Proteins and Nedd8-MLN4924 Adduct by Mass Spectrometry ». *Cell Biochemistry and Biophysics* 67(1): 139-47.
- Ye, Yihong, et Michael Rape. 2009. « Building ubiquitin chains: E2 enzymes at work ». *Nature reviews. Molecular cell biology* 10(11): 755-64.
- Yin, Qing et al. 2019. « K27-linked ubiquitination of BRAF by ITCH engages cytokine response to maintain MEK-ERK signaling ». *Nature Communications* 10. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6478693/> (12 juillet 2019).
- Yuan, W., et R. M. Krug. 2001. « Influenza B Virus NS1 Protein Inhibits Conjugation of the Interferon (IFN)-Induced Ubiquitin-like ISG15 Protein ». *The EMBO journal* 20(3): 362-71.
- Yuan, Wei-Chien et al. 2014. « K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking ». *Molecular Cell* 54(4): 586-600.
- Zeng, Ming et al. 2016. « CRL4 Wdr70 Regulates H2B Monoubiquitination and Facilitates Exo1-Dependent Resection ». *Nature Communications* 7(1): 1-11.
- Zhang, Wenjing et al. 2008. « Genetic Analysis of CAND1-CUL1 Interactions in Arabidopsis Supports a Role for CAND1-Mediated Cycling of the SCFTIR1 Complex ». *Proceedings of the National Academy of Sciences of the United States of America* 105(24): 8470-75.
- Zhang, Yinghua et al. 2012. « Transcriptional Regulation of the Ufm1 Conjugation System in Response to Disturbance of the Endoplasmic Reticulum Homeostasis and Inhibition of Vesicle Trafficking ». *PLOS ONE* 7(11): e48587.

- Zhao, Ran et al. 2014. « DNA Damage-Binding Complex Recruits HDAC1 to Repress Bcl-2 Transcription in Human Ovarian Cancer Cells ». *Molecular cancer research: MCR* 12(3): 370-80.
- Zheng, Ning et al. 2002. « Structure of the Cul1-Rbx1-Skp1-F BoxSkp2 SCF Ubiquitin Ligase Complex ». *Nature* 416(6882): 703-9.
- Zou, Yongxin et al. 2009. « Characterization of Nuclear Localization Signal in the N Terminus of CUL4B and Its Essential Role in Cyclin E Degradation and Cell Cycle Progression ». *The Journal of Biological Chemistry* 284(48): 33320-32.